A wavelet analysis of the Rosenblatt process: chaos expansion and estimation of the self-similarity parameter

Jean-Marc Bardet, Ciprian A. Tudor

A wavelet analysis of the Rosenblatt process: chaos expansion and estimation of the self-similarity parameter

J.-M. Bardet[a,*], C.A. Tudor †, [b]

[a]*SAMM, Université de Paris 1, 90, rue de Tolbiac, 75634, Paris, France*
[b]*Laboratoire Paul Painlevé, Université de Lille 1, F-59655 Villeneuve d'Ascq, France.*

Abstract

By using chaos expansion into multiple stochastic integrals, we make a wavelet analysis of two self-similar stochastic processes: the fractional Brownian motion and the Rosenblatt process. We study the asymptotic behavior of the statistic based on the wavelet coefficients of these processes. Basically, when applied to a non-Gaussian process (such as the Rosenblatt process) this statistic satisfies a non-central limit theorem even when we increase the number of vanishing moments of the wavelet function. We apply our limit theorems to construct estimators for the self-similarity index and we illustrate our results by simulations.

Keywords: multiple Wiener-Itô integral, wavelet analysis, Rosenblatt process, fractional Brownian motion, noncentral limit theorem, self-similarity, parameter estimation
2000 MSC: Primary: 60G18, Secondary 60F05, 60H05, 62F12

1. Introduction

The self-similarity property for a stochastic process means that scaling of time is equivalent to an appropriate scaling of space. That is, a process $(Y_t)_{t\geq 0}$ is self-similar of order $H>0$ if for all $c>0$ the processes $(Y_{ct})_{t\geq 0}$ and $(c^H Y_t)_{t\geq 0}$ have the same finite dimensional distributions. This property is crucial in applications such as network traffic analysis, mathematical finance, astrophysics, hydrology or image processing. We refer to the monographs [7], [14] or [28] for complete expositions on theoretical and practical aspects of self-similar stochastic processes.

The most popular self-similar process is the fractional Brownian motion (fBm). Its practical applications are numerous. This process is defined as a centered Gaussian process $(B_t^H)_{t\geq 0}$ with covariance function

$$R_H(t,s) = \mathbb{E}(B_t^H B_s^H) = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right), \qquad t,s\geq 0. \tag{1}$$

It can be also defined as the only Gaussian self-similar process with stationary increments. Recently, this stochastic process has been widely studied from the stochastic calculus point of view as well as from the statistical analysis point of view. Various types of stochastic integrals with respect to it have been introduced and several types of stochastic differential equations driven by fBm have been considered. Other stochastic processes which are self-similar with stationary increments are the Hermite processes (see [8], [12], [29]); an Hermite process of order q is actually an iterated integral of a deterministic function with q variables with respect to the standard Brownian motion. These processes appear as limits in the so-called noncentral limit theorem and they have the same covariance as the fBm. The fBm is obtained for $q=1$ and it is the only Gaussian Hermite process. For $q=2$ the corresponding process is known as the Rosenblatt process. Although it received a less important attention than the fractional Brownian motion, this process is still of interest in practical applications because of its self-similarity, stationarity and long-range dependence of increments.

*Corresponding author
Email addresses: bardet@univ-paris1.fr (J.-M. Bardet), tudor@math.univ-lille1.fr (C.A. Tudor †,)
† Associate member: SAMM, Université de Paris 1, 90, rue de Tolbiac, 75634, Paris, France

Actually the great popularity of the fractional Brownian motion in practice (hydrology, telecommunications) is due to these properties; one prefers fBm rather than higher order Hermite process because it is a Gaussian process and the corresponding stochastic analysis is much easier. But in concrete situations when empirical data attest to the presence of self-similarity and long memory without the Gaussian property (an example is mentioned in the paper [30]), one can use a Hermite process living in a higher chaos, in particular the Rosenblatt process.

When studying self-similar processes, a question of major interest is to estimate their self-similarity order. This is important because the self-similarity order characterizes in some sense the process: for example in the fBm case as well as for Hermite processes this order gives the long-range dependence property of its increments and it characterizes the regularity of the trajectories. Several statistics, applied directly to the process or to its increments, have been introduced to address the problem of estimating the self-similarity index. Naturally, parametric statistics (exact or Whittle approached maximum likelihood) estimators have been considered. But to enlarge the method to a more general class of models (such as e.g. locally or asymptotically self-similar models), it could be interesting to apply semiparametric methods such as wavelets based, log-variogram or log-periodogram estimators. Information and details on these various approaches can be found in the books of Beran [7] and Doukhan *et al.* [13].

Our purpose is to develop a wavelet-based analysis of the fBm and the Rosenblatt process by using multiple Wiener-Itô integrals and to apply our results to estimate the self-similarity index. More precisely, let $\psi : \mathbb{R} \to \mathbb{R}$ be a continuous function with support included in the interval $[0, 1]$ (called "mother wavelet"). Assume that there exists an integer $Q \geq 1$ such that

$$\int_{\mathbb{R}} t^p \psi(t) dt = 0 \text{ for } p = 0, 1, \ldots, Q-1 \tag{2}$$

and

$$\int_{\mathbb{R}} t^Q \psi(t) dt \neq 0.$$

The integer $Q \geq 1$ is called the *number of vanishing moments.* For a stochastic process $(X_t)_{t\in[0,N]}$ and for a "scale" $a \in \mathbb{N}^*$ we define its wavelet coefficient by

$$d(a, i) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \psi\left(\frac{t}{a} - i\right) X_t dt = \sqrt{a} \int_0^1 \psi(x) X_{a(x+i)} dx \tag{3}$$

for $i = 1, 2, \ldots, N_a$ with $N_a = [N/a] - 1$. Let us also introduce the normalized wavelet coefficient

$$\tilde{d}(a, i) = \frac{d(a, i)}{(\mathbb{E}d^2(a, i))^{\frac{1}{2}}} \tag{4}$$

and the statistic

$$V_N(a) = \frac{1}{N_a} \sum_{i=1}^{N_a} \left(\tilde{d}^2(a, i) - 1\right). \tag{5}$$

The wavelet analysis consists in the study of the asymptotic behavior of the sequence $V_N(a)$ when $N \to \infty$. Usually, if X is a stationary long-memory process or a self-similar second-order process, then $\mathbb{E}d^2(a, i)$ is a power-law function of a with exponent $2H - 1$ (when $a \to \infty$) or $2H + 1$ respectively. Therefore, if $V_N(a)$ converges to 0, a log-log-regression of $\frac{1}{N_a} \sum_{i=1}^{N_a} d^2(a_j, i)$ onto a_j will provide an estimator of H (with an appropriate choice of $(a_j)_j$). Hence, the asymptotic behavior of $V_N(a)$ will completely give the asymptotic behavior of the estimator (see the Section 5 for details). There are four main advantages to use such an estimator based on wavelets: firstly, it is a semiparametric method that can be easily generalized. Secondly, it is based on the log-regression of the wavelet coefficient sample variances onto several scales and the graph of such a regression provides interesting information concerning the goodness-of-fit of the model (χ^2 goodness-of-fit have been defined and studied in [3] or [4] from this log-regression). Thirdly, it is a computationally efficient estimator (this is due to the Mallat's algorithm for computing the wavelet coefficients). Finally, it is a very robust method: it is not sensitive to possible polynomial trends as soon as the number of vanishing moments Q is large enough.

This method has been introduced by Flandrin [15] in the case of fBm and later it has been extended to more general processes in [1]. The asymptotic behavior of the estimator obtained from the wavelet variation has been specified in the case of long-memory Gaussian processes in [6] or [21], for long memory linear processes in [27] and for multiscale fractional Gaussian processes in [4]. However the case of the Rosenblatt process has not yet been studied (note that the wavelet synthesis of Rosenblatt processes was treated in [2] and we will use this method in the section devoted to simulations).

In our paper we will use a recently developed theory based on Malliavin calculus and Wiener-Itô multiple stochastic integrals. Let us briefly recall these new results. In [25] the authors gave necessary and sufficient conditions for a sequence of random variables in a fixed Wiener chaos (that means in essence that these random variables are iterated integrals of a fixed order with respect to a given Brownian motion) to converge to a standard Gaussian random variable (one of these conditions is that the sequence of the fourth order moments converges to 3 which represents the moment of order 4 of a standard Gaussian random variable). Another equivalent condition is given in the paper [24] in terms of the Malliavin derivative. These results created a powerful link between the Malliavin calculus and limit theorems and they have already been used in several papers (for example in [33], [9] and [10] to study the variations of the Hermite processes).

Recall (see [15] and [3]) that if $X = B^H$ is a fBm in (5) then the following fact happens: for any $Q > 1$ and $H \in (0,1)$ the statistic $V_N(a)$ renormalized by $\sqrt{N}$ converges to a centered Gaussian random variable. If $Q = 1$ then the barrier $H = 3/4$ appears: that is, the behavior of $V_N(a)$ is normal (*i.e.* it satisfies a central limit theorem) only if $H \in (0,3/4)$ and we prove in Section 3 that the limit of $V_N(a)$ (normalized by N^{2-2H}) is a Rosenblatt random variable R_1^H when $H \in (3/4,1)$. This non-central limit theorem is new in the case of the wavelet based statistic (it is known for other type of statistics, such as the statistic constructed from the variations of the process, see [33]). In this case we also prove that the limit holds only in law and not in $\mathbb{L}^2$ in contrast to the case of quadratic variations studied in [33].

The study of $V_N(a)$ in the Rosenblatt process case (see Section 3, formula (17) for the definition of the Rosenblatt process) with $H \in (1/2,1)$ put in light interesting and somehow intriguing phenomena. *The main fact is that the number of vanishing moments Q does not affect its convergence and the limit of $V_N(a)$ is always non-Gaussian* (it is still Rosenblatt). This fact is unexpected and different to the situations known in the literature. Actually, the statistic V_N can be decomposed into two parts: a term in the fourth chaos (an iterated integral of order 4 with respect to a Wiener process) and a term in the second chaos. We analyze here both terms and we deduce that the term in the fourth Wiener chaos keeps some of the characteristics of the Gaussian case (it has to be renormalized by $\sqrt{N}$ and it has a Gaussian limit for $H \in (1/2,3/4)$). But the main term in $V_N(a)$ which gives the normalization is the second chaos term and its detailed analysis shows that the normalization depends on H (it is of order of N^{1-H}) and its limit is (in law) a Rosenblatt random variable.

The consequences of these results are also interesting for the wavelet based estimator of the self-similarity order. Assume that a sample $(X_1, \cdots, X_N)$ is obseved, where X is a fBm or a Rosenblatt process. First, by approximating the wavelet coefficients (3), we consider a statistic $\widehat{V}_N(a)$ that can be computed from the observations $(X_1, \cdots, X_N)$ and we prove that the limit theorems satisfied by $V_N(a)$ also hold for $\widehat{V}_N(a)$ as soon as a is large enough with respect to N. Secondly, we deduce the convergence rates for the wavelet based estimator of H following the cases: $Q \geq 2$ and X is a fBm, $Q = 1$, $H \in (3/4,1)$ and X is a fBm or $H \in (\frac{1}{2},1)$ and X is a Rosenblatt process. The regularity of ψ also plays a role. For practical use, it is clear that if X is a fBm it is better to chose $Q \geq 2$ and the mother wavelet ψ to be twice continuously differentiable (for example, this is the case when ψ is a Daubechies wavelet with order ≥ 8). On the other hand, if X

is a Rosenblatt process the number of vanishing moments Q plays no role. Our simulations illustrate the convergence of $\widehat{V}_N(a)$ and of the estimator of H in this last case.

Our results open other related questions: for a stationary long-memory Gaussian or linear process, parametric estimators (such as the Whittle's estimator, see [17] and [18]) or semiparametric estimators (such as wavelet based estimator, see [21] and [27]) satisfy a central limit theorem with an "usual" convergence rate (that is $\sqrt{N}$ for Whittle's estimator and $\sqrt{N/a}$ for wavelet based estimator). This is not the case of the Whittle's estimator for non-linear functionals of a long-memory Gaussian process (see [19]). As we prove in our paper, this is no longer the case for the wavelet estimator based on the observation of a Rosenblatt process. Therefore it seems that the Rosenblatt processes and the second order polynomials of Gaussian process give the same asymptotic behavior for the estimators. Indeed, in [19] it has been established that the Whittle's estimator in the case of second order polynomials of Gaussian processes satisfies a noncentral limit theorem with a Rosenblatt distribution as a limit and with the same convergence rate as in Theorem 4 below. Several interesting questions arise naturally from this: firstly, are there any reciprocal results true? (*i.e.* may we suspect that the Whittle's estimator of the long-memory parameter for the increments of the Rosenblatt process and the wavelet based estimator of the long-memory parameter for a second order polynomial of a long-memory Gaussian process also satisfy the same noncentral limit theorem?). Secondly, are the conclusions the same for other semiparametric estimators such as log-periodogram or local Whittle's estimators?

We organized the paper as follows. Section 2 contains some preliminaries on multiple Wiener-Itô integrals with respect to the Brownian motion. In Section 3 we treat the situation when the driven process is the fBm. In this case our new result is the noncentral limit theorem satisfied by the wavelet based statistic proved in Theorem 2. In Section 4 we enter into a non-Gaussian world: our observed process is the Rosenblatt process and using the techniques of the Malliavin calculus and recent interesting results for the convergence of sequence of multiple stochastic integrals, we study in details the sequence $V_N(a)$. In Section 5 we construct an observable estimator based on the approximated wavelet coefficients and we study its asymptotic behavior. We also construct an estimator of the self-similarity order and we illustrate its convergence by numerical results.

2. Preliminaries

2.1. Basic tools on multiple Wiener-Itô integrals

Let $(W_t)_{t\in[0,T]}$ be a classical Wiener process on a standard Wiener space $(\Omega,\mathcal{F},P)$. If $f\in\mathbb{L}^2([0,T]^n)$ with $n\geq 1$ integer, we introduce the multiple Wiener-Itô integral of f with respect to W. We refer to [23] for a detailed exposition of the construction and the properties of multiple Wiener-Itô integrals.

Let $f\in\mathcal{S}_n$, that means that there exists $n\geq 1$ integer such that

$$f=\sum_{i_1,\ldots,i_n} c_{i_1,\ldots,i_n}1_{A_{i_i}\times\ldots\times A_{i_n}}$$

where the coefficients satisfy $c_{i_1,\ldots,i_n}=0$ if two indices i_k and i_ℓ are equal and the sets $A_i\in\mathcal{B}([0,T])$ are disjoint. For such a step function f we define

$$I_n(f)=\sum_{i_1,\ldots,i_n} c_{i_1,\ldots,i_n}W(A_{i_1})\ldots W(A_{i_n})$$

where we set $W([a,b])=W_b-W_a$. It can be seen that the application I_n constructed above from $\mathcal{S}_n$ equipped with the scaled norm $\frac{1}{\sqrt{n!}}\|\cdot\|_{\mathbb{L}^2([0,T]^n)}$ to $\mathbb{L}^2(\Omega)$ is an isometry on $\mathcal{S}_n$, i.e. for m,n positive integers,

$$\begin{aligned}\mathbb{E}\left(I_n(f)I_m(g)\right) &= n!\langle f,g\rangle_{\mathbb{L}^2([0,T]^n)} \quad \text{if } m=n, &(6)\\ \mathbb{E}\left(I_n(f)I_m(g)\right) &= 0 \quad \text{if } m\neq n.\end{aligned}$$

It also holds that

$$I_n(f) = I_n(\tilde{f}) \tag{7}$$

where $\tilde{f}$ denotes the symmetrization of f defined by $\tilde{f}(x_1,\dots,x_n) = \frac{1}{n!}\sum_{\sigma\in\mathcal{S}_n} f(x_{\sigma(1)},\dots,x_{\sigma(n)})$.
Since the set $\mathcal{S}_n$ is dense in $\mathbb{L}^2([0,T]^n)$ for every $n\geq 1$, the mapping I_n can be extended to an isometry from $\mathbb{L}^2([0,T]^n)$ to $\mathbb{L}^2(\Omega)$ and the above properties (6) and (7) hold true for this extension. Note also that I_n can be viewed as an iterated stochastic integral

$$I_n(f) = n!\int_0^1\int_0^{t_n}\dots\int_0^{t_2} f(t_1,\dots,t_n)dW_{t_1}\dots dW_{t_n}.$$

We recall the product formula for two multiple integrals (see [23]): if $f\in\mathbb{L}^2([0,T]^n)$ and $g\in\mathbb{L}^2([0,T]^m)$ are symmetric functions, then it holds that

$$I_n(f)I_m(g) = \sum_{\ell=0}^{m\wedge n} \ell! C_m^\ell C_n^\ell I_{m+n-2\ell}(f\otimes_\ell g) \tag{8}$$

where $C_m^\ell = \frac{m!}{\ell!(m-\ell)!}$ and the contraction $f\otimes_\ell g$ belongs to $\mathbb{L}^2([0,T]^{m+n-2\ell})$ for $\ell = 0,1,\dots,m\wedge n$ with

$$(f\otimes_\ell g)(s_1,\dots,s_{n-\ell},t_1,\dots,t_{m-\ell}) = \int_{[0,T]^\ell} f(s_1,\dots,s_{n-\ell},u_1,\dots,u_\ell)g(t_1,\dots,t_{m-\ell},u_1,\dots,u_\ell)du_1\dots du_\ell.$$

In general the contraction $f\otimes_\ell g$ is not necessarily a symmetric function even if the two functions f and g are symmetric.

3. The case of the fractional Brownian motion

3.1. A presentation using chaos expansion

We will assume in this part that $X = B^H$ a (normalized) fractional Brownian motion (fBm in the sequel) with Hurst parameter $H\in(0,1)$. Recall that B^H is a centered Gaussian process with covariance function (1). It is the only normalized Gaussian H-self-similar process with stationary increments. Recall also the fBm $(B_t^H)_{t\in[0,N]}$ with Hurst parameter $H\in(0,1)$ can be written

$$B_t^H = \int_0^t K^H(t,s)dW_s, \quad t\in[0,N]$$

where $(W_t, t\in[0,N])$ is a standard Wiener process and for $s<t$, and for $H>\frac{1}{2}$ the kernel K^H has the expression

$$K^H(t,s) = c_H s^{\frac{1}{2}-H}\int_s^t (u-s)^{H-\frac{3}{2}}u^{H-\frac{1}{2}}du \tag{9}$$

with $c_H = \left(\frac{H(2H-1)}{\beta(2-2H,H-\frac{1}{2})}\right)^{\frac{1}{2}}$ and $\beta(\cdot,\cdot)$ the beta function. For $t>s$ and for every $H\in(0,1)$ the partial derivative of K^H with respect to its first variable is given by

$$\frac{\partial K^H}{\partial t}(t,s) = \partial_1 K^H(t,s) = c_H\left(\frac{s}{t}\right)^{\frac{1}{2}-H}(t-s)^{H-\frac{3}{2}}. \tag{10}$$

In this case it is trivial to decompose in chaos the wavelet coefficient $d(a,i)$. By a stochastic Fubini theorem we can write

$$\begin{aligned} d(a,i) &= \sqrt{a}\int_0^1 \psi(x)B_{a(x+i)}^H dx = \sqrt{a}\int_0^1\psi(x)dx\left(\int_0^{a(x+i)} dB_u^H\right)\\ &= \sqrt{a}\int_0^1\psi(x)dx\int_0^{a(x+i)} K^H(a(x+i),u)dW_u = I_1\left(f_{a,i}(\cdot)\right)\end{aligned}$$

where I_1 denotes the multiple integral of order one (actually, it is the Wiener integral with respect to W) and

$$f_{a,i}(u) = 1_{[0,a(i+1)]}(u)\sqrt{a}\int_{(\frac{u}{a}-i)\vee 0}^{1} \psi(x)K^H(a(x+i),u)dx. \tag{11}$$

Thus, for all $a > 0$ and $i \in \mathbb{N}$,

$$\mathbb{E}\big(d^2(a,i)\big) = \|f_{a,i}\|^2_{\mathbb{L}^2([0,N])} = a^{2H+1}C_\psi(H)$$
$$\text{with} \quad C_\psi(H) = -\frac{1}{2}\int_0^1\int_0^1 \psi(x)\psi(x')|x-x'|^{2H}dx\,dx' \tag{12}$$

(see [3]). This formula is essential for the estimation of H (see Section 5). Using the product formula (8)

$$I_1(f)I_1(g) = I_2(f\otimes g) + \langle f,g\rangle_{\mathbb{L}^2([0,N])}$$

we obtain

$$V_N(a) = \frac{1}{N_a}\sum_{i=1}^{N_a}\Big(\frac{I_2\big(f_{a,i}^{\otimes 2}\big) + \|f_{a,i}\|^2_{\mathbb{L}^2([0,N])}}{(\mathbb{E}d(a,i))^2} - 1\Big) = I_2(f_N^{(a)})$$

where

$$f_N^{(a)} = a^{-2H-1}C_\psi(H)^{-1}\frac{1}{N_a}\sum_{i=1}^{N_a} f_{a,i}^{\otimes 2}. \tag{13}$$

As a consequence the wavelet statistic V_N can be expressed as a multiple Wiener-Itô integral of order two and therefore the chaos expansion techniques can be applied to its analysis.

3.2. A multidimensional Central Limit Theorem satisfied by $\big(V_N(a_i)\big)_{1\le i\le \ell}$.

When the observed process is a fBm with $H < 3/4$, the statistic $V_N(a)$ satisfies a central limit theorem. This fact is known and we will not insist on this case. We just recall it to situate it in our context. Since $\mathbb{E}I_2^2(f) = 2!\|f\|^2_{\mathbb{L}^2[0,N]^{\otimes 2}}$, if $(a_i)_{1\le i\le \ell}$ is a family of integer numbers such that $a_i = i\,a$ for $i = 1,\dots,\ell$ and $a \in \mathbb{N}^*$, it holds that

$$\begin{aligned}
\mathrm{Cov}\,(V_N(a_p),V_N(a_q)) &= 2!(p\,q\,a^2)^{-2H-1}C_\psi(H)^{-2}\frac{1}{N_{a_p}}\frac{1}{N_{a_q}}\sum_{j=1}^{N_{a_p}}\sum_{j'=1}^{N_{a_q}}\langle f_{a_p,j}^{\otimes 2}, f_{a_q,j'}^{\otimes 2}\rangle_{\mathbb{L}^2([0,N])^{\otimes 2}}\\
&= 2(p\,q\,a^2)^{-2H-1}C_\psi(H)^{-2}\frac{1}{N_{a_p}}\frac{1}{N_{a_q}}\sum_{j=1}^{N_{a_p}}\sum_{j'=1}^{N_{a_q}}\langle f_{a_p,j}, f_{a_q,j'}\rangle^2_{\mathbb{L}^2([0,N])}.
\end{aligned}$$

We know (see e.g. [3] and [4]) that:

$$\begin{aligned}
\langle f_{a_p,j}, f_{a_q,j'}\rangle_{\mathbb{L}^2([0,N])} &= \mathbb{E}\,(d(a_p,j)d(a_q,j'))\\
&= -\frac{1}{2}(p\,q\,a^2)^{1/2}a^{2H}\int_0^1\int_0^1 \psi(x)\psi(x')|px-qx'+pj-qj'|^{2H}dx\,dx'.
\end{aligned} \tag{14}$$

and using the Taylor expansion and the property (2) satisfied by ψ,

$$\langle f_{a_p,j}, f_{a_q,j'}\rangle^2_{\mathbb{L}^2([0,N])} = p\,q\,a^{4H+2}\mathcal{O}\big(1+|pj-qj'|\big)^{4H-4Q}$$
$$\Longrightarrow \quad |\mathrm{Cov}\,(V_N(a_p),V_N(a_q))| \le C\,\frac{1}{N_{a_q}^2}\sum_{j=1}^{N_{a_p}}\sum_{j'=1}^{N_{a_q}}\mathcal{O}\big(1+|pj-qj'|\big)^{4H-4Q}.$$

Consequently, if $Q > 1$ and $H \in (0,1)$ or if $Q = 1$ and $H \in (0,3/4)$,

$$N_a\,\mathrm{Cov}\,(V_N(a_p),V_N(a_q)) \underset{N\to\infty}{\longrightarrow} \ell_1(p,q,H) \quad \text{with}$$
$$\ell_1(p,q,H) = \frac{1}{2\,d_{pq}\,(p\,q)^{2H-1}}\sum_{k=-\infty}^{\infty}\Big(\frac{1}{C_\psi(H)}\int_0^1\int_0^1 \psi(x)\psi(x')|px-qx'+kd_{pq}|^{2H}dx\,dx'\Big)^2, \tag{15}$$

where d_{pq} is a constant depending only on p and q. Moreover, the following limit theorem holds true.

Theorem 1. *Let $(X_t)_{t\geq 0}$ be a fBm, $V_N(a)$ be defined by (5) and $L_1^{(\ell)}(H) = (\ell_1(p,q,H))_{1\leq p,q\leq \ell}$ with $\ell_1(p,q,H)$ defined in (15). Then if $Q>1$ and $H\in(0,1)$ or if $Q=1$ and $H\in(0,3/4)$, for all $a>0$,*

$$\sqrt{N_a}\big(V_N(i\,a)\big)_{1\leq i\leq \ell} \xrightarrow[N_a\to\infty]{\mathcal{D}} \mathcal{N}_m\big(0\,,\,L_1^{(\ell)}(H)\big). \tag{16}$$

Proof: It is well-known in the literature (see e.g. [3]). ■

The above central limit theorem does not hold for $Q=1$ and $H\in(\frac{3}{4},1)$. We will focus on this case in the following paragraph.

3.3. A noncentral limit theorem satisfied by $V_N(a)$.

We need at this point to define the Rosenblatt process. The Rosenblatt process with time interval $[0,N]$, denoted in the sequel by $(R_t^H)_{t\in[0,N]}$ appears as a limit in the so-called *noncentral limit theorem* (see [12] or [29]). It is not a Gaussian process and can be defined through its representation as a double iterated integral with respect to a standard Wiener process (see [32]). More precisely, the Rosenblatt process with self-similarity order $H\in(\frac{1}{2},1)$ is defined by

$$R_t^H = \int_0^t\int_0^t L_t^H(y_1,y_2)dW_{y_1}dW_{y_2} \tag{17}$$

where $(W_t, t\in[0,N])$ is a standard Brownian motion and

$$L_t^H(y_1,y_2) = d_H 1_{[0,t]}(y_1)1_{[0,t]}(y_2)\int_{y_1\vee y_2}^t \partial_1 K^{H'}(u,y_1)\partial_1 K^{H'}(u,y_2)du, \tag{18}$$

with K^H the standard kernel defined in (9) appearing in the Wiener integral representation of the fBm, its derivatives being defined in (10) and

$$H' = \frac{H+1}{2} \quad \text{and} \quad d_H = \frac{1}{H+1}\left(\frac{H}{2(2H-1)}\right)^{-\frac{1}{2}}. \tag{19}$$

The random variable R_1^H is called a Rosenblatt random variable with self-similarity index H. A Rosenblatt process is a process having stationary increments and

- it is H-self-similar in the sense that for any $c>0$, $(R_{ct}^H)_{t\geq 0} =^{(d)} (c^H R_t^H)_{t\geq 0}$, where " $=^{(d)}$ " means equivalence of all finite dimensional distributions;
- $\mathbb{E}(|R_t^H|^p)<\infty$ for any $p>0$, and R^H has the same variance and covariance as a standard fractional Brownian motion with parameter H.
- the paths of the Rosenblatt process are Hölder continuous of order $\delta<H$.

We obtain the following noncentral limit theorem for the wavelet coefficient of the fBm with $H>\frac{3}{4}$. Define

$$\ell_2(H) \quad = \quad \Big(\frac{2H^2(2H-1)}{4H-3}\Big)^{1/2}\frac{\big(\int_0^1 x\,\psi(x)\,dx\big)^2}{C_\psi(H)} = \frac{1}{d_{2H-1}}\frac{\big(\int_0^1 x\,\psi(x)\,dx\big)^2}{C_\psi(H)}. \tag{20}$$

Then,

Theorem 2. *Let $(X_t)_{t\geq 0}$ be a fBm and $V_N(a)$ be defined by (5). If $Q=1$ and $\frac{3}{4}<H<1$,*

$$\ell_2^{-1}(H)\,N_a^{2-2H}V_N(a) \xrightarrow[N_a\to\infty]{\mathcal{D}} R_1^{H_0}.$$

where $H_0=2H-1$, $\ell_2(H)$ is defined by (20) and $R_1^{H_0}$ is a Rosenblatt random variable given by (17).

Proof: With $f_N^{(a)}$ defined as in (13), we can write

$$N_a^{2-2H} V_N(a) = N_a^{2-2H}\, I_2\big(f_N^{(a)}\big).$$

Using the expression (11) of the kernel $f_{a,i}$, one has

$$\begin{aligned} f_N^{(a)}(y_1,y_2) = \frac{1}{a^{2H} C_\psi(H)} \frac{1}{N_a} \sum_{i=1}^{N_a} 1_{[0,a(i+1)]}(y_1) 1_{[0,a(i+1)]}(y_2) \\ \times \int_{(\frac{y_1}{a}-i)\vee 0}^{1} \int_{(\frac{y_2}{a}-i)\vee 0}^{1} \psi(x)\psi(z) K^H\left(a(x+i),y_1\right) K^H\left(a(z+i),y_2\right) dxdz. \end{aligned}$$

To show that the sequence $\ell_2^{-1}(H) N_a^{2-2H} I_2\big(f_N^{(a)}\big)$ converges in law to the Rosenblatt random variable R_1^{2H-1} it suffices to show that its cumulants converge to the cumulants of R_1^{2H-1} (see e.g. [17] where it has been proven that the law of a multiple integral of order 2 is given by its cumulants). On the other hand, we know (see [17], [25]) that the k-cumulant of a random variable $I_2(f)$ in the second Wiener chaos can be computed as follows

$$c_k(I_2(f)) = \int_{[0,1]^k} dy_1 \dots dy_k f(y_1,y_2) f(y_2,y_3) \dots f(y_{k-1},y_k) f(y_k,y_1). \tag{21}$$

In particular the kth cumulant of the Rosenblatt random variable R_1^{2H-1} is given by

$$c_k(R_1^{2H-1}) = d_{2H-1}^k (H(2H-1))^k \int_{[0,1]^k} \left[|x_1-x_2|\cdot \ldots \cdot |x_{k-1}-x_k| \cdot |x_k-x_1|\right]^{2H-2} dx_1 \dots dx_k.$$

We compute

$$\begin{aligned} & c_k\big(C_\psi(H) N_a^{2-2H} I_2\big(f_N^{(a)}\big)\big) \\ &= N_a^{(2H-2)k} N_a^{-k} \sum_{i_1,\dots,i_k=1}^{N_a} \int_{[0,1]^k} dy_1 \dots dy_k \int_{[0,1]^{2k}} dx_1 dz_1 \dots dx_k dz_k \psi(x_1)\psi(z_1)\psi(x_2)\psi(z_2)\dots\psi(x_k)\psi(z_k) \\ &\quad \times K^H(a(x_1+i_1),y_1) K^H(a(z_1+i_1),y_2) K^H(a(x_2+i_2),y_2) K^H(a(z_2+i_2),y_3) \times \dots \\ &\quad \dots \times K^H(a(x_{k-1}+i_{k-1}),y_{k-1}) K^H(a(z_k+i_k),y_k) K^H(a(x_k+i_k),y_k) K^H(a(z_k+i_k),y_1). \end{aligned}$$

Using Fubini Theorem and the fact that (for every x, x', i, j, a)

$$\begin{aligned} \int_0^{a(x+i)\wedge a(x'+j)} K^H(a(x+i),y_1) K^H(a(x'+j),y_1) dy_1 &= R_H(a(x+i), a(x'+j)) \\ &= a^{2H} R_H(x+i, x'+j) \end{aligned}$$

(from the representation of the fBm as a Wiener integral with respect to the Wiener process) we get

$$
\begin{aligned}
& c_k\big(C_\psi(H) N_a^{2-2H} I_2\big(f_N^{(a)}\big)\big) \\
= & \; N_a^{(2H-2)k} N_a^{-k} \sum_{i_1,\dots,i_k=1}^{N_a} \int_{[0,1]^{2k}} dx_1 dz_1 \dots dx_k dz_k \psi(x_1)\psi(z_1)\psi(x_2)\psi(z_2)\dots\psi(x_k)\psi(z_k) \\
& R_H(z_1+i_1, x_2+i_2) R_H(z_2+i_2, x_3+i_3)\dots R_H(z_{k-1}+i_{k-1}, x_k+i_k) R_H(z_k+i_k, x_1+i_1) \\
= & \; N_a^{(2H-2)k} a^{2Hk} N_a^{-k} \sum_{i_1,\dots,i_k=1}^{N_a} \int_{[0,1]^{2k}} dx_1 dz_1 \dots dx_k dz_k \psi(x_1)\psi(z_1)\psi(x_2)\psi(z_2)\dots\psi(x_k)\psi(z_k) \\
& \times \big[|z_1 - x_2 + i_1 - i_2| \cdot |z_2 - x_3 + i_2 - i_3| \dots \cdot |z_{k-1} - x_k + i_{k-1} - i_k| \cdot |z_k - x_1 + i_k - i_1|\big]^{2H} \\
= & \; N_a^{(2H-2)k} N_a^{-k} \sum_{i_1,\dots,i_k=1}^{N_a} (|i_1 - i_2| \dots \cdot |i_{k-1} - i_k| \cdot |i_k - i_1|)^{2H} \\
& \times \int_{[0,1]^{2k}} dx_1 dz_1 \dots dx_k dz_k \psi(x_1)\psi(z_1)\psi(x_2)\psi(z_2)\dots\psi(x_k)\psi(z_k) \left| \Big(1 + \frac{z_1 - x_2}{i_1 - i_2})\Big)^{2H} \dots \Big(1 + \frac{z_k - x_1}{i_k - i_1})\Big)^{2H} \right| \\
\sim & \; N_a^{(2H-2)k} H^k (2H-1)^k N_a^{-k} \sum_{i_1,\dots,i_k=1}^{N_a} (|i_1 - i_2| \dots \cdot |i_{k-1} - i_k| \cdot |i_k - i_1|)^{2H-2} \\
& \int_{[0,1]^{2k}} dx_1 dz_1 \dots dx_k dz_k \psi(x_1)\psi(z_1)\psi(x_2)\psi(z_2)\dots\psi(x_k)\psi(z_k) x_1 z_1 \dots x_k z_k
\end{aligned}
$$

and we used the fact that the integral of the mother wavelet vanishes and a Taylor expansion of second order of the function $(1+x)^{2H}$. By $a_n \sim b_n$ we mean that the sequences a_n and b_n have the same limit as $n \to \infty$. As a consequence, by a standard Riemann sum argument, the sequence

$$
N_a^{(2H-2)k} N_a^{-k} \sum_{i_1,\dots,i_k=1}^{N_a} \big(|i_1 - i_2| \times \dots \times |i_{k-1} - i_k||i_k - i_1|\big)^{2H-2} = N_a^{-k} \sum_{i_1,\dots,i_k=1}^{N_a} \Big(\frac{|i_1 - i_2| \times \dots \times |i_{k-1} - i_k||i_k - i_1|}{N_a}\Big)^{2H-2}
$$

converges to the integral

$$
\int_{[0,1]^k} \big(|x_1 - x_2| \times \dots \times |x_{k-1} - x_k||x_k - x_1|\big)^{2H-2} dx_1 \dots dx_k
$$

and therefore it is clear that the cumulant of $\ell_2^{-1}(H) N_a^{2-2H} I_2\big(f_N^{(a)}\big)$ converges to the k cumulant of the Rosenblatt random variable R_1^{2H-1} (see [29], [32]). ■

In the case of the statistic based on the variations of the fBm, in the case $H \in (3/4, 1)$ the statistic $\frac{1}{N} \sum_{i=0}^{N-1} \frac{(B^H_{\frac{i+1}{N}} - B^H_{\frac{i}{N}})^2}{N^{-2H}} - 1$, renormalized by a constant times N^{2-2H}, converges in $\mathbb{L}^2(\Omega)$ to a Rosenblatt random variable at time 1 (see [33]). In the wavelet world, our above result gives only the convergence in law. The following question is then natural: can we get $\mathbb{L}^2$-convergence for the renormalized statistic $V_N(a)$? The answer is negative and a proof of this fact can be found in the extended version of our paper available on `arXiv`. But we will present below a brief and heuristic argument to see what the $\mathbb{L}^2$ convergence does not hold. The term $f_N^{(a)}$ can be written as

$$\begin{aligned} f_N^{(a)}(y_1,y_2) &= \frac{1}{a^{2H}C_\psi(H)}\frac{1}{N_a}\sum_{i=1}^{N_a} 1_{[0,a(i+1)]}(y_1)1_{[0,a(i+1)]}(y_2) \\ &\times\Big(1_{[0,ai]}(y_1)1_{[0,ai]}(y_2)\int_0^1\int_0^1 dxdz\psi(x)\psi(z)K^H(a(x+i),y_1)K^H(a(z+i),y_2) \\ &\quad +1_{[0,ai]}(y_1)1_{[ai,a(1+i)]}(y_2)\int_0^1\int_{\frac{y_2}{a}-i}^1 dxdz\psi(x)\psi(z)K^H(a(x+i),y_1)K^H(a(z+i),y_2) \\ &\quad +1_{[0,ai]}(y_2)1_{[ai,a(1+i)]}(y_1)\int_{\frac{y_1}{a}-i}^1\int_0^1 dxdz\psi(x)\psi(z)K^H(a(x+i),y_1)K^H(a(z+i),y_2) \\ &\quad\quad +1_{[ai,a(1+i)]}(y_1)1_{[ai,a(1+i)]}(y_2)\int_{\frac{y_1}{a}-i}^1\int_{\frac{y_2}{a}-i}^1 dxdz\psi(x)\psi(z)K^H(a(x+i),y_1)K^H(a(z+i),y_2)\Big) \\ &= f_N^{(a,1)}(y_1,y_2)+f_N^{(a,2)}(y_1,y_2)+f_N^{(a,3)}(y_1,y_2)+f_N^{(a,4)}(y_1,y_2). \end{aligned}$$

It can be shown that the terms $N_a^{2-2H}f_N^{(a,2)}$, $N_a^{2-2H}f_N^{(a,3)}$ and $N_a^{2-2H}f_N^{(a,4)}$ converge to zero in $\mathbb{L}^2([0,\infty)^2)$ as $N_a\to\infty$. It remains to understand the convergence of the term $f_N^{(a,1)}$. Using again the property (2) of the mother wavelet ψ we can write

$$\begin{aligned} f_N^{(a,1)}(y_1,y_2) = \frac{1}{a^{2H}C_\psi(H)}\frac{1}{N_a}\sum_{i=0}^{N_a}\int_0^1\int_0^1 dxdz\psi(x)\psi(z)1_{[0,ai]}(y_1)1_{[0,ai]}(y_2) \\ \times\left(K^H\left(a(x+i),y_1\right)-K^H\left(ai,y_1\right)\right)\left(K^H\left(a(z+i),y_2\right)-K^H\left(ai,y_2\right)\right). \end{aligned}$$

Therefore, with $\alpha(a,i,x)$ and $\beta(a,i,z)$ located in $[ai,ax+ai]$ and $[ai,az+ai]$ respectively,

$$\begin{aligned} I_2\big(f_N^{(a,1)}\big) &= \frac{1}{a^{2H}C_\psi(H)}\frac{1}{N_a}I_2\Big(\sum_{i=0}^{N_a}\int_0^1\int_0^1 dxdz\psi(x)\psi(z)1_{[0,ai]}(y_1)1_{[0,ai]}(y_2) \\ &\qquad \times ax\,\partial_1K^H(\alpha(a,i,x),y_1)\times az\,\partial_1K^H(\beta(a,i,z),y_2)\Big). \end{aligned}$$

By approximating the points the points $\alpha(a,i,x)$ and $\beta(a,i,z)$ by ai and since (from an usual approximation of a sum by a Riemann integral) when $N_a\to\infty$, with $y_1,y_2\in[0,N]$,

$$\begin{aligned} \sum_{i=0}^{N_a}1_{[0,ai]}(y_1)1_{[0,ai]}(y_2)\partial_1K^H(ai,y_1)\partial_1K^H(ai,y_2) &\sim \int_{(y_1\vee y_2)/a}^{N_a}du\,\partial_1K^H(au,y_1)\partial_1K^H(au,y_2) \\ &\sim \frac{1}{d_{2H-1}\,a}\,L_N^{2H-1}(y_1,y_2) \end{aligned}$$

where L_N^{2H-1} is the kernel of the Rosenblatt process with self-similarity index $2H-1$ (see its definition in (18)), the sequence $\ell_2^{-1}(H)N_a^{2-2H}f_N^{(a,1)}$ is equivalent (in the sense that it has the same limit pointwise) to $N^{1-2H}L_N^{2H-1}$. Then, in some sense, $\ell_2^{-1}(H)N_a^{2-2H}V_N(a)$ is equivalent to $N^{1-2H}I_2(L_N^{2H-1}) = N^{1-2H}R_N^{2H-1}=Z_1^{2H-1}$ but this equivalence is only in law. The fact that the sequence $N_a^{2-2H}V_N(a)$ is not Cauchy in $\mathbb{L}^2$ comes from the fact that the sequence $N^{1-2H}R_N^{2H-1}$ is not Cauchy in $\mathbb{L}^2$ as it can be easily seen by computing the square mean of the difference $N^{1-2H}R_N^{2H-1}-M^{1-2H}R_M^{2H-1}$.

4. The Rosenblatt case

4.1. Chaotic expansion of the wavelet variation

We study in this section the limit of the wavelet-based statistic V_N given by (5) in the situation when the observed process X is a Rosenblatt process. Throughout this section, assume that $X=R^H$ is a Rosenblatt process with self-similarity order $H>1/2$ defined via the stochastic integral representation (17). We first express the statistic $V_N(a)$ in terms of multiple stochastic integrals. The wavelet coefficient of the Rosenblatt process can be written as

$$\begin{aligned} d(a,i) &= \sqrt{a}\int_0^1 \psi(x) R^H_{a(x+i)}dx \\ &= \sqrt{a}\int_0^1 \psi(x)dx\Big(\int_0^{a(x+i)}\int_0^{a(x+i)} L^H_{a(x+i)}(y_1,y_2)dW_{y_1}dW_{y_2}\Big) \\ &= I_2\left(g_{a,i}(\cdot,\cdot)\right) \end{aligned}$$

where, with H' and d_H defined in (19),

$$g_{a,i}(y_1,y_2) = d_H\,\sqrt{a}\,1_{[0,a(i+1)]}(y_1)1_{[0,a(i+1)]}(y_2) \\ \times \int_{(\frac{y_1\vee y_2}{a}-i)\vee 0}^1 dx\,\psi(x)\,\Big(\int_{y_1\vee y_2}^{a(x+i)} \partial_1 K^{H'}(u,y_1)\partial_1 K^{H'}(u,y_2)du\Big)$$

for every $y_1,y_2\geq 0$. The product formula for multiple stochastic integrals (8) gives

$$I_2(f)I_2(g) = I_4(f\otimes g) + 4I_2(f\otimes_1 g) + 2\langle f,g\rangle_{\mathbb{L}^2(0,N]^{\otimes 2}}$$

if $f,g\in\mathbb{L}^2([0,N]^2)$ are two symmetric functions and the contraction $f\otimes_1 g$ is defined by

$$(f\otimes_1 g)(y_1,y_2) = \int_0^N f(y_1,x)g(y_2,x)dx.$$

Thus, we obtain

$$d^2(a,i) \;=\; I_4\left(g_{a,i}^{\otimes 2}\right) + 4I_2\left(g_{a,i}\otimes_1 g_{a,i}\right) + 2\|g_{a,i}\|^2_{\mathbb{L}^2[0,N]^2}$$

and note that, since the covariance of the Rosenblatt process is the same as the covariance of the fractional Brownian motion, we will also have

$$\mathbb{E}\big(d^2(a,i)\big) = \mathbb{E}\left(I_2(g_{a,i})\right)^2 = 2\|g_{a,i}\|^2_{\mathbb{L}^2[0,N]^2} = a^{2H+1}C_\psi(H).$$

Therefore, we obtain the following decomposition for the statistic $V_N(a)$:

$$V_N(a) \;=\; a^{-2H-1}C_\psi(H)^{-1}\frac{1}{N_a}\left[\sum_{i=1}^{N_a} I_4\left(g_{a,i}^{\otimes 2}\right) + 4\sum_{i=1}^{N_a} I_2\left(g_{a,i}\otimes_1 g_{a,i}\right)\right] = T_2+T_4$$

$$\text{with}\quad \left\{\begin{array}{lcl} T_2 &=& a^{-2H-1}C_\psi(H)^{-1}\frac{4}{N_a}\sum_{i=1}^{N_a} I_2\left(g_{a,i}\otimes_1 g_{a,i}\right) \\ T_4 &=& a^{-2H-1}C_\psi(H)^{-1}\frac{1}{N_a}\sum_{i=1}^{N_a} I_4\left(g_{a,i}^{\otimes 2}\right)\end{array}\right. . \tag{22}$$

To understand the limit of the sequence V_N we need to regard the two terms above (note that similar terms appear in the decomposition of the variation statistic of the Rosenblatt process, see [33]). In essence, the following will happen: the term T_4 which lives in the fourth Wiener chaos keeps some characteristics of the fBm case (since it has to be renormalized by $\sqrt{N_a}$ except in the case $Q=1$ and $H>\frac{3}{4}$ where the normalization is N_a^{2-2H}) and its limit will be Gaussian (except for $Q=1$ and $H>\frac{3}{4}$). Unfortunately, this somehow nice behavior does not affect the limit of V_N which is the same as the limit of T_2 and therefore it is non-normal (see below).

Now, let us study the asymptotic behavior of $\mathbb{E}T_4^2$. From (22), we have

$$T_4 = I_4(g_N^{(a)}) \quad \text{where} \quad g_N^{(a)} = a^{-2H-1}C_\psi(H)^{-1}\frac{1}{N_a}\sum_{i=1}^{N_a} g_{a,i}^{\otimes 2}, \tag{23}$$

and thus, by the isometry of multiple stochastic integrals,

$$\mathbb{E}T_4^2 = 4!\|\tilde{g}_N^{(a)}\|^2_{\mathbb{L}^2[0,N]^4}$$

where by $\tilde{g}_N^{(a)}$ we denoted the symmetrization of the function $g_N^{(a)}$ is its four variables. Since $\|\tilde{g}_N^{(a)}\|^2_{\mathbb{L}^2[0,N]^4} \leq \|g_N^{(a)}\|^2_{\mathbb{L}^2[0,N]^4}$ we obtain

$$\begin{aligned}
\mathbb{E}T_4^2 &\leq 4!\, C_\psi(H)^{-2} a^{-4H-2} \frac{1}{N_a^2} \sum_{i,j=1}^{N_a} \langle g_{a,i}^{\otimes 2}, g_{a,j}^{\otimes 2}\rangle_{\mathbb{L}^2[0,N]^4} \\
&\leq 4!\, C_\psi(H)^{-2} a^{-4H-2} \frac{1}{N_a^2} \sum_{i,j=1}^{N_a} \langle g_{a,i}, g_{a,j}\rangle^2_{\mathbb{L}^2(0,N]^{\otimes 2}}.
\end{aligned}$$

But,

$$\langle g_{a,i}, g_{a,j}\rangle_{\mathbb{L}^2(0,N]^{\otimes 2}} = \frac{1}{2}\,\mathbb{E}(d(a,i)d(a,j)).$$

Again, using the fact the fBm and the Rosenblatt process have the same covariance, we obtain the same behavior (up to a multiplicative constant) as in the case of the fractional Brownian motion. That is, using (15) and the proof of Theorem 2

Proposition 1. *Let $(X_t)_t = (R_t^H)_{t\geq 0}$ be a Rosenblatt process and T_4 be defined by (22). If $Q > 1$ and $H \in (\frac{1}{2}, 1)$ or if $Q = 1$ and $H \in (\frac{1}{2}, \frac{3}{4})$, then with $\ell_1(1,1,H)$ defined in (15)*

$$N_a \mathbb{E}(T_4^2) \leq N_a 4!\|g_N^{(a)}\|_{\mathbb{L}^2[0,N]^4} \underset{N\to\infty}{\longrightarrow} 3\,\ell_1(1,1,H) \tag{24}$$

and, if $Q = 1$ and $H \in (\frac{3}{4}, 1)$ then with $\ell_2(H)$ defined in (20),

$$N_a^{4-4H}\,\mathbb{E}(T_4^2) \leq N_a^{4-4H} 4!\|g_N^{(a)}\|_{\mathbb{L}^2[0,N]^4} \underset{N\to\infty}{\longrightarrow} 3\,\ell_2(H). \tag{25}$$

The above result gives only an upper bound for the L^2 norm of the term T_4; this will be sufficient to obtain the desired limit of the sequence $V_N(a)$ in the Rosenblatt case. It will follow from the following paragraph that the L^2 norm of the term T_4 will be dominated by the L^2 norm of the term T_2 and therefore the limit of T_2 will be the limit of the statistics $V_N(a)$.

4.2. Asymptotic behavior of the term T_2

We study here the term in second Wiener chaos that appears in the decomposition of $V_N(a)$. We have

$$T_2 = I_2(h_N^{(a)}) \quad \text{with} \quad h_N^{(a)} = 4\,\frac{1}{a^{2H+1}C_\psi(H)}\frac{1}{N_a}\sum_{i=1}^{N_a} g_{a,i}\otimes_1 g_{a,i}. \tag{26}$$

We first compute the contraction $g_{a,i}\otimes_1 g_{a,i}$. We have

$$\begin{aligned}
(g_{a,i}\otimes_1 g_{a,i})(y_1,y_2) &= \int_0^N g_{a,i}(y_1,z)g_{a,i}(y_2,z)dz \\
&= a\,d_H^2\,1_{[0,a(i+1)]}(y_1)1_{[0,a(i+1)]}(y_2)\int_0^{a(i+1)} dz\Big[\int_{(\frac{y_1\vee z}{a}-i)\vee 0}^1 dx\,\psi(x)\Big(\int_{y_1\vee z}^{a(x+i)}\partial_1 K^{H'}(u,y_1)\partial_1 K^{H'}(u,z)du\Big)\Big] \\
&\quad\times\Big[\int_{(\frac{y_2\vee z}{a}-i)\vee 0}^1 dx'\,\psi(x')\Big(\int_{y_2\vee z}^{a(x'+i)}\partial_1 K^{H'}(u',y_2)\partial_1 K^{H'}(u',z)du'\Big)\Big] \\
&= a\,d_H^2\,1_{[0,a(i+1)]}(y_1)1_{[0,a(i+1)]}(y_2)\Big(\Big[\int_{(\frac{y_1}{a}-i)\vee 0}^1 dx\,\psi(x)\int_{(\frac{y_2}{a}-i)\vee 0}^1 dx'\,\psi(x') \\
&\quad\times\int_{y_1}^{a(x+i)}\int_{y_2}^{a(x'+i)} M(u,y_1,u',y_2)\,du\,du'\int_0^{u\wedge u'} M(u,z,u',z)dz\Big]
\end{aligned}$$

where $M(u,y_1,u',y_2)=\partial_1 K^{H'}(u,y_1)\partial_1 K^{H'}(u',y_2)$ and $H'=(H+1)/2$. Now, we have already seen that $\int_0^{t\wedge s} K^H(t,z)K^H(s,z)dz = R_H(t,s)$ with $R_H(t,s)$ given in (1) and therefore (see [23], Chapter 5)

$$\int_0^{u\wedge u'} M(u,z,u',z)\,dz = H'(2H'-1)\,|u-u'|^{2H'-2} \tag{27}$$

(In fact, this relation can be easily derived from $\int_0^{u\wedge v} K^{H'}(u,y_1)K^{H'}(v,y_1)dy_1 = R_{H'}(u,v)$, and will be used repeatedly in the sequel). Thus denoting $\alpha_H = H'(2H'-1) = H(H+1)/2$ and since ψ is $[0,1]$-supported, we obtain

$$\begin{aligned}(g_{a,i}\otimes_1 g_{a,i})(y_1,y_2) \quad &= \quad a\,d_H^2\alpha_H\,1_{[0,a(i+1)]}(y_1)1_{[0,a(i+1)]}(y_2)\int_{\left(\frac{y_1}{a}-i\right)\vee 0}^1\int_{\left(\frac{y_2}{a}-i\right)\vee 0}^1 dxdx'\psi(x)\psi(x')\\ &\times\int_{y_1}^{a(x+i)}\int_{y_2}^{a(x'+i)}|u-u'|^{2H'-2}M(u,y_1,u',y_2)\,du\,du'.\end{aligned}$$

By direct computation, it is possible to evaluate the expectation of T_2^2 as follows (the proof can be found on the extended version on `arXiv`)

$$\begin{aligned}N_a^{2-2H}\mathbb{E}T_2^2 \;\underset{N\to\infty}{\longrightarrow}\; & 32\,\frac{\alpha_H^4 d_H^4}{H(2H-1)C_\psi^2(H)}\Big(\int_{[0,1]^4}\psi(x)\psi(x')xx'|ux-vx'|^{2H'-2}dxdx'dudv\Big)^2\\ \underset{N\to\infty}{\longrightarrow}\; & C_{T_2}^2(H)=\frac{32(2H-1)}{H(H+1)^2}\Big(\frac{C_\psi(H')}{C_\psi(H)}\Big)^2=\Big(4d(H)\frac{C_\psi(H')}{C_\psi(H)}\Big)^2,\end{aligned} \tag{28}$$

where we used $\int_{[0,1]^2} xx'|ux-vx'|^{2H'-2}dudv = \int_0^x\int_0^{x'}|u'-v'|^{2H'-2}du'dv' = \frac{1}{2H'(2H'-1)}\big(|x|^{2H'}+|x'|^{2H'}-|x-x'|^{2H'}\big)$ and $\alpha_H^4\,d_H^4 = \frac14(2H-1)^2H^2$. We do not prove here this estimate because it is a consequence of the following proposition which show that the sequence $C_{T_2}^{-1}(H)N_a^{1-H}T_2$ (and therefore the sequence $V_N(a)$) converges in distribution to a Rosenblatt random variable with self-similarity index H.

Proposition 2. *Let $(R_t^H)_{t\geq 0}$ be a Rosenblatt process and let T_2 be the sequence given by (22) and computed from $(R_t^H)_{t\geq0}$. Then, for any $Q\geq 1$ and $H\in(\frac12,1)$, with C_{T_2} given by (28), there exists a Rosenblatt random variable R_1^H with self-similarity order H such as*

$$C_{T_2}^{-1}(H)\,N_a^{1-H}T_2 \;\underset{N\to\infty}{\overset{\mathcal{D}}{\longrightarrow}}\; R_1^H.$$

Proof: This proof follows the lines of the proof of Theorem 2. With $T_2 = I_2\big(h_N^{(a)}\big)$ in mind, as in the proof of Theorem 2, a direct proof that cumulants of the sequence $N_a^{1-H}I_2\big(h_N^{(a)}\big)$ converge to those of the Rosenblatt process can be given. Indeed, since the random variable $N_a^{1-H}I_2\big(h_N^{(a)}\big)$ is an element of the second Wiener chaos, its cumulants can be computed by using formula (21). By using the key formula (27)

$$\begin{aligned}&c_k\big(N_a^{1-H}I_2\big(h_N^{(a)}\big)\big)\\ &= N_a^{k(1-H)}\int_{[0,1]^k}dy_1\dots dy_k h_N^{(a)}(y_1,y_2)h_N^{(a)}(y_2,y_3)\times\cdots\times h_N^{(a)}(y_k,y_1)\\ &= N_a^{-Hk}a^{-2Hk}C_\psi(H)^{-k}4^k d_H^{2k}\alpha_H^{2k}\sum_{i_1,\dots,i_k=1}^{N_a}\int_{[0,1]^{2k}}\prod_{j=1}^k\psi(x_j)\psi(x'_j)dx_jdx'_j\\ &\times\int_0^{a(x_1+i_1)}\int_0^{a(x'_1+i_1)}\cdots\int_0^{a(x_j+i_j)}\int_0^{a(x'_j+i_j)}du_1du'_1\dots du_jdu'_j\Big(\prod_{j=1}^k|u_j-u'_j|\Big)^{2H'-2}\Big(\prod_{j=1}^k|u'_j-u'_{j+1}|\Big)^{2H'-2}\end{aligned}$$

with the convention $u_{k+1} = u_1$. Next, we will make the change of variables $\tilde{u}_j = au_j$ and $\tilde{u'}_j = au'_j$ and this will simplify the factors containing a. We thus get

$$
\begin{aligned}
&c_k\big(N_a^{1-H} I_2\big(h_N^{(a)}\big)\big) \\
&= N_a^{-Hk} C_\psi(H)^{-k} 4^k d_H^{2k} \alpha_H^{2k} \sum_{i_1,\dots,i_k=1}^{N_a} \int_{[0,1]^{2k}} \prod_{j=1}^{k} \psi(x_j)\psi(x'_j) dx_j dx'_j \\
&\times \int_0^{x_1+i_1}\int_0^{x'_1+i_1}\dots\int_0^{x_k+i_k}\int_0^{x'_k+i_k} du_1 du'_1 \dots du_k du'_k \Big(\prod_{j=1}^{k} |u_j - u'_j|\Big)^{2H'-2} \Big(\prod_{j=1}^{k} |u'_j - u'_{j+1}|\Big)^{2H'-2}.
\end{aligned}
$$

We will note at this point that the vanishing moment property of the wavelet function ψ allows us to replace integration intervals $[0, x_j + i_j]$ by intervals $[i_j, x_j + i_j]$. After doing this, we perform the change of variables $\tilde{u}_j = u_j - i_j$, $\tilde{u'}_j = u'_j - i_j$ (for every $j = 1, \dots, k$) to obtain

$$
\begin{aligned}
&c_k\big(N_a^{1-H} I_2\big(h_N^{(a)}\big)\big) \\
&= N_a^{-Hk} C_\psi(H)^{-k} 4^k d_H^{2k} \alpha_H^{2k} \sum_{i_1,\dots,i_k=1}^{N_a} \int_{[0,1]^{2k}} \prod_{j=1}^{k} \psi(x_j)\psi(x'_j) dx_j dx'_j \\
&\times \int_0^{x_1}\int_0^{x'_1}\dots\int_0^{x_k}\int_0^{x'_k} du_1 du'_1 \dots du_k du'_k \Big(\prod_{j=1}^{k} |u_j - u'_j|\Big)^{2H'-2} \Big(\prod_{j=1}^{k} |u'_j - u'_{j+1} + i_j - i_{j+1}|\Big)^{2H'-2},
\end{aligned}
$$

and then, with changes of variables $\tilde{u}_j = \frac{u_j}{x_j}$ and $\tilde{u'}_j = \frac{u'_j}{x_j}$, we can write

$$
\begin{aligned}
&c_k\big(N_a^{1-H} I_2\big(h_N^{(a)}\big)\big) \\
&= N_a^{-Hk} C_\psi(H)^{-k} 4^k d_H^{2k} \alpha_H^{2k} \sum_{i_1,\dots,i_k=1}^{N_a} \int_{[0,1]^{2k}} \prod_{j=1}^{k} x_j x'_j\, \psi(x_j)\psi(x'_j) dx_j dx'_j \\
&\times \int_{[0,1]^{2k}} du_1 du'_1 \dots du_k du'_k \Big(\prod_{j=1}^{k} |x_j u_j - x'_j u'_j|\Big)^{2H'-2} \Big(\prod_{j=1}^{k} |x_j u'_j - x'_j u'_{j+1} + i_j - i_{j+1}|\Big)^{2H'-2} \\
&= N_a^{-Hk} C_\psi(H)^{-k} 4^k d_H^{2k} \alpha_H^{2k} \sum_{i_1,\dots,i_k=1}^{N_a} \int_{[0,1]^{2k}} \prod_{j=1}^{k} x_j x'_j\, \psi(x_j)\psi(x'_j) dx_j dx'_j \\
&\times \sum_{i_1,\dots,i_k=1}^{N_a} \prod_{j=1}^{k} |i_j - i_{j+1}|^{2H'-2} \int_{[0,1]^{2k}} du_1 du'_1 \dots du_k du'_k \Big(\prod_{j=1}^{k} |u_j x_j - u'_j x'_j|\Big)^{2H'-2} \prod_{j=1}^{k} \Big|1 + \frac{u'_j x'_j - u_{j+1} x_{j+1}}{i_j - i_{j+1}}\Big|^{2H'-2}.
\end{aligned}
$$

An analysis of the function $(1+x)^{2H'-2}$ in the vicinity of the origin shows that the cumulant $c_k\big(N_a^{1-H} I_2\big(h_N^{(a)}\big)\big)$ behaves as (we recall that by $a_n \sim b_n$ we mean that the sequences a_n and b_n have the same limit as $n \to \infty$)

$$
\begin{aligned}
c_k\big(N_a^{1-H} I_2\big(h_N^{(a)}\big)\big) &\sim N_a^{-Hk} C_\psi(H)^{-k} 4^k d_H^{2k} \alpha_H^{2k} \int_{[0,1]^{2k}} \prod_{j=1}^{k} \psi(x_j)\psi(x'_j) dx_j dx'_j (x_j x'_j)^k \\
&\times \sum_{i_1,\dots,i_k=1}^{N_a} \prod_{j=1}^{k} |i_j - i_{j+1}|^{2H'-2} \int_{[0,1]^{2k}} du_1 du'_1 \dots du_k du'_k \Big(\prod_{j=1}^{k} |u_j x_j - u'_j x'_j|\Big)^{2H'-2}.
\end{aligned}
$$

By using an usual Riemann sum convergence, we notice that

$$
N_a^{-Hk} \sum_{i_1,\dots,i_k=1}^{N_a} \prod_{j=1}^{k} |i_j - i_{j+1}|^{2H'-2} = N_a^{-k} \sum_{i_1,\dots,i_k=1}^{N_a} \prod_{j=1}^{k} \Big(\frac{|i_j - i_{j+1}|}{N_a}\Big)^{2H'-2}
$$

converges as $N_a \to \infty$ to the integral

$$
\int_{[0,1]^k} dx_1 \dots dx_k \big(|x_1 - x_2| \cdot |x_2 - x_3| \times \dots \times |x_k - x_1|\big)^{2H'-2}
$$

which is $d_H^{-k}\alpha_H^{-k}c_k(R_1^H)$ (here $c_k(R_1^H)$ denotes the cumulant of the random variable R_1^H). We conclude that

$$\begin{aligned}
c_k\big(N_a^{1-H}I_2\big(h_N^{(a)}\big)\big) \quad \underset{N_a\to\infty}{\longrightarrow} \quad & C_\psi(H)^{-k}4^k d_H^k c_k(R_1^H)\int_{[0,1]^{2k}}\prod_{j=1}^k x_j x'_j\psi(x_j)\psi(x'_j)dx_j dx'_j \\
& \times\int_{[0,1]^{2k}} du_1du'_1\dots du_k du'_k\Big(\prod_{j=1}^k |u_jx_j-u'_jx'_j|\Big)^{2H'-2} \\
= \quad & C_\psi(H)^{-k}4^k d_H^k\alpha_H^k\left(\int_0^1\int_0^1\psi(x)\psi(x')\frac{1}{2}(x^{2H'}+(x')^{2H'}-|x-x'|^{2H'}dxdx'\right)^k c_k(R_1^H) \\
= \quad & C_\psi(H)^{-k}4^k d_H^k C_\psi(H')^k\, c_k(R_1^H)
\end{aligned}$$

where we recall the notation $C_\psi(H)=-\frac{1}{2}\int_0^1\int_0^1\psi(x)\psi(x')|x-x'|^{2H'}dxdx'$. As a consequence $C_{T_2}^{-1}T_2$ converges in law to R_1^H where $C_{T_2}=4d_H\frac{C_\psi(H')}{C_\psi(H)}$. ■

We finally state our main result on the convergence of the wavelet statistic constructed from a Rosenblatt process. Its proof is a consequence of Proposition 1 and 2.

Theorem 3. *Let $(X_t)_{t\ge0}=(R_t^H)_{t\ge0}$ be a Rosenblatt process and $V_N(a)$ be defined by (5). Then, for any $Q\ge1$ and $H\in(\frac{1}{2},1)$, there exists a Rosenblatt random variable R_1^H with self-similarity order H such as*

$$C_{T_2}^{-1}(H)\,N_a^{1-H}V_N(a)\underset{N\to\infty}{\overset{\mathcal{D}}{\longrightarrow}}R_1^H,$$

where $C_{T_2}(H)$ is given by (28).

It is also possible to provide a multidimensional counterpart of Theorem 3 in the case of a vector of scales $(ai)_{1\le i\le\ell}$ where $\ell\in\mathbb{N}^*$.

Theorem 4. *Let $(X_t)_{t\ge0}=(R_t^H)_{t\ge0}$ be a Rosenblatt process and $V_N(a)$ be defined by (5). Then for every $Q\ge1$ and $H\in(\frac{1}{2},1)$ it holds that*

$$\Big(\frac{N_{ai}^{1-H}}{C_{T_2}(H)}\,V_N(ai)\Big)_{1\le i\le\ell}\underset{N\to\infty}{\overset{\mathcal{D}}{\longrightarrow}}\big(R_{1,1}^H,\dots,R_{1,\ell}^H\big)$$

where $R_{1,i}^H$ are normalized Rosenblatt random variables for all $i=1,\cdots,\ell$ and for all $\lambda_1,\dots,\lambda_\ell\in\mathbb{R}$ the k-th cumulant of the random variable $\sum_{j=1}^\ell\lambda_jR_{1,j}^H$ is

$$\sum_{j_1,\dots,j_k=1}^{\ell}\lambda_{j_1}\dots\lambda_{j_k}c_k(R_1^H)$$

where $c_k(R_1^H)$ denotes the k-th cumulant of the Rosenblatt random variable with self-similarity order H.

Remark 1. *Notice that, since the components of the vector $\Big(R_{1,1}^H,\dots,R_{1,\ell}^H\Big)$ are random variables in the second Wiener case, its finite dimensional distributions are completely determined by the cumulants. Moreover, we deduce that the asymptotic covariance matrix of the random vector $\big(N_a^{1-H}\,V_N(ai)\big)_{1\le i\le\ell}$ is*

$$\Sigma_\ell=C_{T_2}^2(H)\,\big((i\,j)^{1-H}\big)_{1\le i,j\le\ell}. \tag{29}$$

Proof: The proof follows the lines of the proof of Theorem 3. Since for every $\lambda_1,\dots,\lambda_m\in\mathbb{R}$ the linear combination $U_\ell:=\sum_{j=1}^\ell\lambda_j\frac{N_{aj}^{1-H}}{C_{T_2}(H)}V_N(aj)$ is a multiple integral of order two, it is possible to compute its cumulants by using the formula (21). We will obtain the following expression for the k-th cumulant of U_ℓ,

$$\begin{aligned}
c_k(U_\ell)=&\Big(\frac{N_a^{1-H}}{C_{T_2}(H)}\Big)^k\int_{[0,1]^k}dy_1\dots dy_k\big(\lambda_1h_N^{(a)}(y_1,y_2)+\dots+\frac{\lambda_\ell}{\ell^{1-H}}h_N^{(\ell a)}(y_1,y_2)\big) \\
&\times\dots\times\big(\lambda_1h_N^{(a)}(y_k,y_1)+\dots+\frac{\lambda_\ell}{\ell^{1-H}}h_N^{(\ell a)}(y_k,y_1)\big)
\end{aligned}$$

and then, as in the proof of Proposition 2, we can write

$$c_k(U_\ell) = \Big(\frac{4d_H^2\alpha_H^2 a^{-2H} N_a^{-H}}{C_\psi(H)C_{T_2}(H)}\Big)^k \sum_{j_1,\dots,j_k=1}^{\ell} \frac{\lambda_{j_1}\dots\lambda_{j_k}}{(j_1\dots j_k)^H} \sum_{i_1=1}^{N_{aj_1}}\dots\sum_{i_k=1}^{N_{aj_k}} \int_{[0,1]^{2k}} dx_1dx_1'\dots dx_kdx_k' \prod_{q=1}^k \psi(x_q)\psi(x_q')$$
$$\int_0^{aj_1(x_1+i_1)} du_1 \int_0^{aj_1(x_1'+i_1)} du_1'\dots \int_0^{aj_k(x_k+i_k)} du_k \int_0^{aj_k(x_k'+i_k)} du_k' \Big(\prod_{l=1}^k |u_l-u_l'||u_l'-u_{l+1}|\Big)^{2H'-2}$$

where we used again the convention $u_{k+1}=u_1$. We proceed as previously in the proof of Proposition 2 (with changes of variable) and then we obtain

$$c_k(U_\ell) = \Big(\frac{4d_H^2\alpha_H^2 N_a^{-H}}{C_\psi(H)C_{T_2}(H)}\Big)^k \sum_{j_1,\dots,j_k=1}^{\ell} \frac{\lambda_{j_1}\dots\lambda_{j_k}}{(j_1\dots j_k)^{-1}} \sum_{i_1=1}^{N_{aj_1}}\dots\sum_{i_k=1}^{N_{aj_k}} \int_{[0,1]^{2k}} dx_1dx_1'\dots dx_kdx_k' \prod_{q=1}^k x_qx_q'\psi(x_q)\psi(x_q')$$
$$\int_{[0,1]^{2k}} du_1du_1'\dots du_kdu_k' \Big(\prod_{l=1}^k |u_lx_l-u_l'x_l'||u_l'x_l'j_l-u_{l+1}x_{l+1}j_{l+1}+i_lj_l-i_{l+1}j_{l+1}|\Big)^{2H'-2}.$$

Thus, when $N_a\to\infty$,

$$\begin{aligned}
c_k(U_\ell) \quad &\sim \quad \Big(\frac{4d_H^2\alpha_H^2 N_a^{-H}}{C_\psi(H)C_{T_2}(H)}\Big)^k \sum_{j_1,\dots,j_k=1}^{\ell} \frac{\lambda_{j_1}\dots\lambda_{j_k}}{(j_1\dots j_k)^{-1}} \sum_{i_1=1}^{N_{aj_1}}\dots\sum_{i_k=1}^{N_{aj_k}} \int_{[0,1]^{2k}} dx_1dx_1'\dots dx_kdx_k' \prod_{q=1}^k x_qx_q'\psi(x_q)\psi(x_q') \\
&\qquad \Big(\prod_{l=1}^k |i_lj_l-i_{l+1}j_{l+1}|^{2H'-2}\Big) \int_{[0,1]^{2k}} du_1du_1'\dots du_kdu_k' \Big(\prod_{l=1}^k |u_lx_l-u_l'x_l'|\Big)^{2H'-2} \\
&\sim \quad \Big(\frac{4d_H^2\alpha_H\, a^{2H}\, C_\psi(H')\, N_a^{-H}}{C_\psi(H)C_{T_2}(H)}\Big)^k \sum_{j_1,\dots,j_k=1}^{\ell} \frac{\lambda_{j_1}\dots\lambda_{j_k}}{(j_1\dots j_k)^{-1}} \sum_{i_1=1}^{N_{aj_1}}\dots\sum_{i_k=1}^{N_{aj_k}} \Big(\prod_{l=1}^k |i_lj_l-i_{l+1}j_{l+1}|^{2H'-2}\Big) \\
&\underset{N_a\to\infty}{\longrightarrow} \quad \Big(\frac{4d_H^2\alpha_H\, a^{2H}\, C_\psi(H')}{C_\psi(H)C_{T_2}(H)}\Big)^k \sum_{j_1,\dots,j_k=1}^{\ell} \lambda_{j_1}\dots\lambda_{j_k} \int_{[0,1]^k} \Big(\prod_{l=1}^k |y_l-y_{l+1}|^{2H'-2}\Big) dy_1\dots dy_k \\
&\underset{N_a\to\infty}{\longrightarrow} \quad \sum_{j_1,\dots,j_k=1}^{\ell} \lambda_{j_1}\dots\lambda_{j_k}\, c_k(R_1^H),
\end{aligned}$$

where R_1^H is a Rosenblatt random variable with index H and $c_k(R_1^H)$ is its k-th cumulant. ∎

Remark 2. *It is possible and instructive to study the behavior of the term T_4 in the cases $Q>1$ and $H\in(\frac12,1)$ or $Q=1$ and $H\in(\frac12,\frac34)$. It can be already seen from its asymptotic variance that it is very close to the Gaussian case. We can actually show this term converges in law to a Gaussian random variable. This fact does not influence the limit of the statistic V_N but we find that it is interesting from a theoretical point of view. We will denote by $C_{T_4}(H)$ a positive constant such that*

$$N_a\mathbb{E}T_4^2 \to_{N_a\to\infty} C_{T_4}^2(H).$$

Then the following holds: suppose that R^H is a Rosenblatt process with self-similarity order H. Suppose that $Q>1$ or $Q=1$ and $H\in(\frac12,\frac34)$. Then with T_4 defined in (22),

$$\sqrt{N_a}\ T_4 \ \underset{N\to\infty}{\overset{\mathcal{D}}{\longrightarrow}}\ \mathcal{N}\big(0\,,\,C_{T_4}^2(H)\big).$$

The proof can be done by using a criteria in [24] in terms of the Malliavin derivatives. In the other case, i.e. $Q=1$ and $H\in(\frac34,1)$, the limit in law of the renormalized V_N is a Rosenblatt random variable.

5. Applications and simulations

Throughout this section $X^H=(X_t^H)_{t\in\mathbb{R}_+}$ denotes a fBm with $H\in(0,1)$ or a Rosenblatt process with $H\in(1/2,1)$.

5.1. Asymptotic normality of the sample variance of approximated wavelet coefficients

Here a sample $(X_0^H, X_1^H, \cdots, X_N^H)$ of X^H is supposed to be observed. For any couple (a,b), define the following approximations of the wavelet coefficients $d(a,b)$ and of the normalized wavelet coefficients $\tilde{d}(a,b)$ defined in (3) and (4):

$$e(a,b) = \frac{1}{\sqrt{a}} \sum_{k=1}^{N} X_k^H \psi\big(\frac{k}{a} - b\big) \quad \text{and} \quad \tilde{e}(a,b) = \frac{e(a,b)}{a^{H+1/2} C_\psi^{1/2}(H)}, \tag{30}$$

The above expression are the usual Riemann approximations. Define also for $a > 0$,

$$\widehat{V}_N(a) = \frac{1}{N_a} \sum_{i=1}^{N_a} \big(\tilde{e}^2(a,i) - 1\big). \tag{31}$$

Remark 3. *These approximations of wavelet coefficients and their sample variance can be directly computed from data for all mother wavelet ψ. In the particular case of a multiresolution analysis with orthogonal discrete wavelet transform (that means $a = 2^j$), the very fast Mallat's algorithm (similar to FFT for the Fourier's transform) can also be applied to obtain a different approximation $e_M(2^j,k)$ of $d(2^j,k)$. However, when $j \to \infty$, $e(2^j,k) \simeq e_M(2^j,k)$. Indeed, let ϕ the scaling function of a multiresolution analysis satisfying $\int \phi(t)dt = 1$ and $\int \psi^2(t)dt = 1$. Then it can be established that (see for instance [21]) $e_M(2^j,k) = 2^{-j/2} \int dt\, \psi\big(2^{-j}t - k\big) \sum_{s=1}^N \phi(t-s) X_s = 2^{-j/2} \sum_{s=1}^N X_s \int_I dt\, \psi\big(2^{-j}(t+s) - k\big)\phi(t)$. When $j \to \infty$, since ϕ is compactly supported on I, we obtain $\psi\big(2^{-j}(t+s) - k\big) \sim \psi\big(2^{-j}s - k\big)$ for $s \in I$. Then $e_M(2^j,k) \sim 2^{-j/2} \sum_{s=1}^N X_s \psi\big(2^{-j}s - k\big) \int_I \phi(t)dt = e(2^j,k)$ because $\int \phi(t)dt = 1$; more precisely, a first order expansion gives $e_M(2^j,k) = e(2^j,k) + O_P(2^{-j})$, with an approximation error which is negligible with respect to the approximation error computed in the following Lemma 1. The very low time consuming of the Mallat's algorithm together with a straightforward computation of $e(2^j,k)$ without multiresolution analysis, provides a clear advantage of the wavelet based estimator of the parameter H with respect to the estimators based on a minimization of a criterion (such as maximum likelihood estimators).*

Now the following result can be proved :

Lemma 1. *Assume that $\psi \in \mathcal{C}^m(\mathbb{R})$ with $m \geq 1$ and ψ is $[0,1]$-supported. Let $(a_k)_{k \in \mathbb{N}}$ be a sequence of integer numbers satisfying $N\, a_N^{-1} \underset{N\to\infty}{\longrightarrow} \infty$ and $a_N \underset{N\to\infty}{\longrightarrow} \infty$. Then, for any $Q \geq 1$,*

$$\mathbb{E}\big|\widehat{V}_N(a_N) - V_N(a_N)\big| \leq C\Big(\frac{1}{\sqrt{a}_N} + \frac{N^H}{a_N^{H+m}} + \frac{N^{H-Q/2}}{a_N^{(m-Q)/2+H}} \mathbb{1}_{(2H-Q>-1)}\Big). \tag{32}$$

Proof: First note that,

$$\mathbb{E}\big[(\tilde{e}(a,i) - \tilde{d}(a,i))^2\big] = \frac{-1}{2\,C_\psi(H)}\Big(\int_0^1\!\!\int_0^1 dt dt' \psi(t)\psi(t')|t-t'|^{2H} + \frac{2}{a}\int_0^1 dt \psi(t) \sum_{k'=0}^{a-1} \psi\big(\frac{k'}{a}\big)\Big(|t+i|^{2H} - \big|t - \frac{k'}{a}\big|^{2H}\Big)$$
$$+ \frac{1}{a^2}\sum_{k,k'=0}^{a-1} \psi\big(\frac{k}{a}\big)\psi\big(\frac{k'}{a}\big)\Big(\big|i + \frac{k}{a}\big|^{2H} + \big|i + \frac{k'}{a}\big|^{2H} - \big|\frac{k}{a} - \frac{k'}{a}\big|^{2H}\Big).$$

¿From standard Taylor expansion, if g is supposed to be m times continuously differentiable and $[0,1]$-supported, for all $a > 0$,

$$\Big|\frac{1}{a}\sum_{k=0}^{a-1} g\big(\frac{k}{a}\big) - \int_0^1 g(t)dt\Big| \leq \sup_{t\in[0,1]} |g^{(m)}(t)|\, \frac{1}{a^m}.$$

and there exists C depending only on H, Q and ψ such that $\Big|\int_0^1 \psi(t)(i+t)^{2H}dt\Big| \leq C\big((1+i)^{2H-Q}\big)$. Therefore, there exists C depending only on H, Q, m and ψ such that for all $a > 0$

$$\Big|\frac{1}{a}\sum_{k=0}^{a-1}\psi\big(\frac{k}{a}\big)\Big| \leq \frac{C}{a^m} \quad \text{and} \quad \Big|\frac{1}{a}\sum_{k=0}^{a-1}\psi\big(\frac{k}{a}\big)\big|i + \frac{k}{a}\big|^{2H}\Big| = C\big((1+i)^{2H-Q} + \frac{(1+i)^{2H}}{a^m}\big).$$

Finally, as it was already proved in [3], there exists C depending only on H and ψ such that for all $m \geq 1$ and $a > 0$,

$$\begin{aligned}
\Big|\frac{1}{a^2}\sum_{k,k'=0}^{a-1}\psi\big(\frac{k}{a}\big)\psi\big(\frac{k'}{a}\big)\big|\frac{k}{a}-\frac{k'}{a}\big|^{2H}-\int_0^1\!\!\int_0^1 dt dt'\psi(t)\psi(t')|t-t'|^{2H}\Big| &\leq \frac{C}{a} \\
\Big|\frac{1}{a}\int_0^1 dt\psi(t)\sum_{k'=0}^{a-1}\psi\big(\frac{k'}{a}\big)\big|t-\frac{k'}{a}\big|^{2H}-\int_0^1\!\!\int_0^1 dt dt'\psi(t)\psi(t')|t-t'|^{2H}\Big| &\leq \frac{C}{a}.
\end{aligned}$$

All those inequalities imply that there exists C depending only on H, Q, m and ψ such that for all $a > 0$,

$$\mathbb{E}\big[(\tilde{e}(a,i)-\tilde{d}(a,i))^2\big] \leq C\,\Big(\frac{1}{a}+\frac{(1+i)^{2H-Q}}{a^m}+\frac{(1+i)^{2H}}{a^{2m}}\Big). \tag{33}$$

Using Cauchy-Schwarz's inequality,

$$\begin{aligned}
\mathbb{E}\big|\widehat{V}_N(a)-V_N(a)\big| &\leq \frac{1}{N_a}\sum_{i=1}^{N_a}\mathbb{E}\big|\tilde{e}^2(a,i)-\tilde{d}^2(a,i)\big| \\
&\leq \frac{1}{N_a}\sum_{i=1}^{N_a}\sqrt{\mathbb{E}\big[\big(\tilde{e}(a,i)-\tilde{d}(a,i)\big)^2\big]}\sqrt{\mathbb{E}\big[\big(\tilde{e}(a,i)+\tilde{d}(a,i)\big)^2\big]} \\
&\leq \frac{1}{N_a}\sum_{i=1}^{N_a}\sqrt{\mathbb{E}\big[\big(\tilde{e}(a,i)-\tilde{d}(a,i)\big)^2\big]}\sqrt{\mathbb{E}\big[8\tilde{d}^2(a,i)+2\big(\tilde{e}(a,i)-\tilde{d}(a,i)\big)^2\big]} \\
&\leq \sqrt{2}\,\Big(\frac{1}{N_a}\sum_{i=1}^{N_a}\mathbb{E}\big[\big(\tilde{e}(a,i)-\tilde{d}(a,i)\big)^2\big]\Big)^{1/2}\Big(\frac{1}{N_a}\sum_{i=1}^{N_a}\mathbb{E}\big[4\tilde{d}^2(a,i)+2\big(\tilde{e}(a,i)-\tilde{d}(a,i)\big)^2\big]\Big)^{1/2} \\
&\leq \sqrt{2}\,\Big(\frac{1}{N_a}\sum_{i=1}^{N_a}\mathbb{E}\big[\big(\tilde{e}(a,i)-\tilde{d}(a,i)\big)^2\big]\Big)^{1/2}\Big(4+2\frac{1}{N_a}\sum_{i=1}^{N_a}\mathbb{E}\big[\big(\tilde{e}(a,i)-\tilde{d}(a,i)\big)^2\big]\Big)^{1/2}
\end{aligned}$$

since $\mathbb{E}\big[\tilde{d}^2(a,i)\big]=1$ by definition. Finally, from inequality (33) and with a large enough,

$$\begin{aligned}
\frac{1}{N_a}\sum_{i=1}^{N_a}\mathbb{E}\big[\big(\tilde{e}(a,i)-\tilde{d}(a,i)\big)^2\big] &\leq C\,\Big(\frac{1}{a}+\frac{\log(N_a)}{N_a\,a^m}\mathbf{1}_{2H-Q\leq -1}+\frac{N_a^{2H-Q}}{a^m}\mathbf{1}_{2H-Q>-1}+\frac{N_a^{2H}}{a^{2m}}\Big) \\
&\leq C\,\Big(\frac{1}{a}+\frac{\log(N)}{N\,a^{m-1}}\mathbf{1}_{2H-Q\leq -1}+\frac{N^{2H-Q}}{a^{m+2H-Q}}\mathbf{1}_{2H-Q>-1}+\frac{N^{2H}}{a^{2(m+H)}}\Big).
\end{aligned}$$

■

We will use Lemma 1 to prove the following result.

Proposition 3. *Assume that $\psi \in \mathcal{C}^m(\mathbb{R})$ with $m \geq 1$ and ψ is $[0,1]$-supported. Let $(a_k)_{k\in\mathbf{N}}$ be a sequence of integer numbers satisfying $N\,a_N^{-1} \underset{N\to\infty}{\longrightarrow} \infty$ and $a_N \underset{N\to\infty}{\longrightarrow} \infty$. Then,*

1. *if X^H is a fBm with $0 < H < 1$ and $Q \geq 2$ or with $0 < H < 3/4$ and $Q = 1$, and if $N\,a_N^{-1-(1\wedge\frac{2m}{3})} \underset{N\to\infty}{\longrightarrow} 0$, then Theorem 1 holds when $V_N(a)$ is replaced by $\widehat{V}_N(a)$.*
2. *if X^H is a fBm with $3/4 < H < 1$ and $Q = 1$, and if $N\,a_N^{-1-(1\wedge\frac{2m}{3})} \underset{N\to\infty}{\longrightarrow} 0$ then Theorem 2 holds when $V_N(a)$ is replaced by $\widehat{V}_N(a)$.*
3. *if X^H is a Rosenblatt process with $1/2 < H < 1$, $Q \geq 1$ and if $N\,a_N^{-2} \underset{N\to\infty}{\longrightarrow} 0$, then Theorems 3 and 4 hold when $V_N(a)$ is replaced by $\widehat{V}_N(a)$.*

Proof of Proposition 3: The three different cases are respectively obtained from the Markov Inequality. Indeed, for $\varepsilon > 0$ and with $\alpha = 1/2,\ 2-2H,\ 1-H$ (respectively),

$$\mathrm{P}\big(N_a^\alpha\big|\widehat{V}_N(a)-V_N(a)\big|>\varepsilon\big) \leq \frac{1}{\varepsilon}\,\frac{N^\alpha}{a^\alpha}\,\mathbb{E}\big|\widehat{V}_N(a)-V_N(a)\big|.$$

Using Lemma 1, it remains to obtain conditions for insuring

$$\frac{N^\alpha}{a^\alpha}\Big(\frac{1}{\sqrt{a}_N}+\frac{N^H}{a_N^{H+m}}+\frac{N^{H-Q/2}}{a_N^{\frac{m-Q}{2}+H}}\mathbf{1}_{2H-Q>-1}\Big)\underset{N\to\infty}{\longrightarrow}0 \tag{34}$$

for all H to show the three cases of Proposition 3.

1. For $\alpha=1/2$, $Q\geq 2$ and $0<H<1$, condition (34) with $\alpha=1/2$ leads to

$$\max\big(N\,a_N^{-2}\,,\,N\,a_N^{-1-\frac{2m}{2H+1}}\,,\,N\,a_N^{-1-\frac{m}{2H+1-Q}}\mathbf{1}_{2H-Q>-1}\big)\leq\max\big(N\,a_N^{-2}\,,\,N\,a_N^{-1-\frac{2m}{3}}\big)\underset{N\to\infty}{\longrightarrow}0.$$

   It induces the condition $N\,a_N^{-1-(1\wedge\frac{2m}{3})}\underset{N\to\infty}{\longrightarrow}0$.
   For $Q=1$ and any $0<H<3/4$, condition (34) leads to

$$\max\big(N\,a_N^{-2}\,,\,N\,a_N^{-1-\frac{2m}{2H+1}}\,,\,N\,a_N^{-1-\frac{m}{2H}}\big)\underset{N\to\infty}{\longrightarrow}0.$$

   Since $0<H<3/4$, it also induces the condition $N\,a_N^{-1-(1\wedge\frac{2m}{3})}\underset{N\to\infty}{\longrightarrow}0$.

2. if $Q=1$ and $3/4<H<1$, then condition (34) with $\alpha=2-2H$ leads to

$$\max\big(N\,a_N^{-1-\frac{1}{4-4H}}\,,\,N\,a_N^{-1-\frac{2m}{4-2H}}\,,\,N\,a_N^{-1-\frac{m}{3-2H}}\mathbf{1}_{Q=1}\,,\,N\,a_N^{-1-\frac{m}{2-2H}}\mathbf{1}_{Q=2}\big)$$
$$\leq\max\big(N\,a_N^{-2}\,,\,N\,a_N^{-1-\frac{2m}{3}}\big)\underset{N\to\infty}{\longrightarrow}0,$$

   and it also induces the condition $N\,a_N^{-1-(1\wedge\frac{2m}{3})}\underset{N\to\infty}{\longrightarrow}0$.

3. if $Q\geq 1$ and $1/2<H<1$, then condition (34) with $\alpha=1-H$ leads to

$$\max\big(N\,a_N^{-1-\frac{1}{2-2H}}\,,\,N\,a_N^{-1-m}\,,\,N\,a_N^{-1-m}\mathbf{1}_{Q=1}\big)\leq\max\big(N\,a_N^{-2}\,,\,N\,a_N^{-1-m}\big)\underset{N\to\infty}{\longrightarrow}0,$$

   which leads to the condition $N\,a_N^{-2}\underset{N\to\infty}{\longrightarrow}0$ since $m\geq 1$. ∎

Remark 4. *For a concrete estimation of H, the conditions between N and a_N provided in Proposition 3 do not depend on H. Usually, when this conditions depend on H, the convergence rate in the model can be improved. An adaptive procedure for estimating the smallest order possible for (a_k) could be also built as in the paper [5]. Anyway, we do not think that conditions provided on $(a_k)_k$ in Proposition 3 are optimal. They could be improved by controlling $\mathbb{E}\big[(\widehat{V}_N(a)-V_N(a))^2\big]$ instead of $\mathbb{E}\big|\widehat{V}_N(a)-V_N(a)\big|$. However, such computations are very long and tedious in the case of the Rosenblatt process (it requires the computations of fourth-order moments) and we have preferred to avoid them. Moreover, the simulations we give below will show that our results are not so far to be optimal.*

Since the case of fBm has been already studied (see for instance [3]) we only provide below the numerical results when X^H is a Rosenblatt process. Thus, we first exhibit the main result of this paper, *i.e.* the limit theorem $C_{T_2}^{-1}(H)\left(\frac{N}{a_N}\right)^{1-H}\widehat{V}_N(a_N)\underset{N\to\infty}{\overset{\mathcal{D}}{\longrightarrow}}R_1^H$ following the procedure described in the sequel.

Concrete procedure of simulations:

- The samples of Rosenblatt processes are obtained following a similar procedure as the one presented in [2]. With more details, to generate a Rosenblatt process sample $(X_j^H)_{1\leq j\leq N}$:
  1. generate a sample of length $1+N*m$ (in practice we use $m=100$) of a fBm with parameter $(H+1)/2$. We use a wavelet based method introduced by Sellan (see [20]) with a Daubechies wavelet of order 10 but a circular matrix embedding method can also be applied (more details can be found in [13]). Note that this sample is normalized and thus $\mathrm{Var}\,[fBm(1)]=1$. Next, one obtains a sample of length $N*m$ of a fractional Gaussian noise (fGn) defined by the increments of the fBm, *i.e.* $fGn(k)=fBm(k+1)-fBm(k)$ with $\mathrm{Var}\,fGn(k)=1$.

H		0.6			0.7			0.8			0.9		
a_N		$[N^{0.4}]$	$[N^{0.5}]$	$[N^{0.6}]$	$[N^{0.4}]$	$[N^{0.5}]$	$[N^{0.6}]$	$[N^{0.4}]$	$[N^{0.5}]$	$[N^{0.6}]$	$[N^{0.4}]$	$[N^{0.5}]$	$[N^{0.6}]$
$N = 500$	ψ_4	3.31	40.0	3.35	1.69	40.8	1.70	1.08	35.3	1.49	1.53	114	1.05
	ψ_{MH}	2.23	2.40	2.11	1.51	1.55	1.80	0.85	0.91	0.94	0.70	0.75	0.89
	ψ_C	2.29	2.45	2.09	1.76	1.57	1.49	1.23	0.85	0.93	1.05	0.74	0.81
$N = 2000$	ψ_4	6772	3.19	2.99	27935	7.2	1.64	37951	8.8	1.56	43888	9.1	0.75
	ψ_{MH}	2.09	1.91	1.99	1.33	1.60	1.40	1.04	1.08	1.30	0.68	0.67	0.77
	ψ_C	1.95	1.81	1.72	1.18	1.41	1.56	1.21	1.01	1.25	0.66	0.72	0.89
$N = 10000$	ψ_4	2.60	2.77	2.86	1.49	1.41	1.43	0.67	0.84	0.93	0.52	0.46	0.59
	ψ_{MH}	1.86	1.79	2.07	1.10	1.24	1.43	0.72	0.70	0.83	0.51	0.54	0.60
	ψ_C	1.55	1.86	1.72	1.01	1.08	1.29	0.64	0.73	0.75	0.51	0.53	0.56

Table 1: *Computation for different choices of ψ, H, N and a_N of $\sqrt{MSE}$ of $C_{T_2}^{-1}(H)\,\big(\frac{N}{a_N}\big)^{1-H}\widehat{V}_N(a_N)$ from 100 independent replications.*

2. Taqqu proved in [29] that: $\big(\frac{1}{n^H}\sum_{i=1}^{[nt]}(fGn^2(i)-1)\big)_{0\le t\le 1} \xrightarrow[n\to\infty]{\mathcal{D}} (R_t^H)_{0\le t\le 1}$, where R^H is a (non-normalized) Rosenblatt process, since $(1+H)/2 > 3/4$ is the parameter of the fGn. Thus, with $n = mN$, $t = j/N$ and $j = 1,\dots,N$, one computes $Y_j = \frac{1}{(mN)^H}\sum_{i=1}^{mj}(fGn^2(i)-1)$. In this way $(Y_j)_{1\le j\le N}$ approximatively provides a path of $(R_{1/N}^H, R_{2/N}^H,\cdots,R_1^H)$.
3. compute now $X_j^H = \sqrt{\frac{2(2H-1)}{H(H+1)^2}}\,N^H\,Y_j$ for $j = 1,\dots,N$: using the H-self similarity of the Rosenblatt process, $(X_1^H,\dots,X_N^H)$ is approximatively a trajectory of a normalized (*i.e.* $\operatorname{Var} R_1^H = 1$) Rosenblatt process.

The Matlab procedures to generate a trajectory of a fBm or a Rosenblatt process can be downloaded from `http://samos.univ-paris1.fr/-Jean-Marc-Bardet`.

- Several mother wavelets ψ are used:
  - The Daubechies' wavelet of order 4, ψ_4 (which is such that $Q = 4$ and $\mathcal{C}^1([0,1])$ but not $\mathcal{C}^2([0,1])$);
  - The Mexican Hat wavelet, ψ_{MH} (which is such that $Q = 2$ and $\psi \in \mathcal{C}^\infty(\mathbb{R})$ and is essentially compactly supported);
  - The function ψ_C such that $\psi_C(t) = t(t-1)(2t-1)(t^2 - t + \frac{1}{7})$ for all $t \in [0,1]$ and $\psi_C = 0$ elsewhere (which is such that $Q = 3$ and $\psi \in \mathcal{C}^\infty([0,1])$ except in 0 and 1).
- The values of constants $C_\psi(H)$, $C_\psi(H')$ and $C_{T_2}(H)$ are obtained from usual approximations of integrals by Riemann sums.

Montecarlo experiments using 100 independent replications of trajectories are realized for each $H = 0.6$, 0.7, 0.8 and 0.9 and for $N = 500$, $N = 2000$ and $N = 10000$. The sequence of scales $(a_N)_N$ is selected to be such that $a_N = [N^{0.4}]$, $a_N = [N^{0.5}]$ or $a_N = [N^{0.6}]$. The following Table 1 provides the results of simulations, which are values $\sqrt{\text{MSE}}$ of $C_{T_2}^{-1}(H)\left(\frac{N}{a_N}\right)^{1-H}\widehat{V}_N(a_N)\big)$ for the different choices of ψ, H, N and a_N.

The main points to highlight from these simulations are the following:

- globally, for N and a_N large enough then $\big(C_{T_2}^{-1}(H)\left(\frac{N}{a_N}\right)^{1-H}\widehat{V}_N(a_N)\big)_N$ seems to converge in distribution to a centered distribution with a variance close to 1;
- if a_N is not large enough (in these simulations, $a_N = [N^{0.4}]$), there is a bias which clearly appears in the case of ψ_4 since this wavelet function is not regular enough. In this way we may see that the conditions required for (a_N) in Proposition 3 are close to be optimal.
- Since Rosenblatt processes are generated from an approximation algorithm based on a wavelet synthesis, the larger H the higher the octave required to obtain a convenient trajectory. Because memories of computers are finite and we expect a low consuming time, the generator of Rosenblatt process has

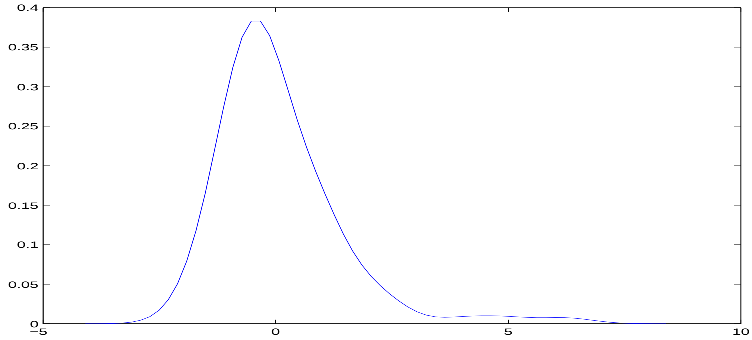

Figure 1: *FFT estimation (Silverman's method) of the density of the limit of* $\left(C_{T_2}^{-1}(H)\left(\frac{N}{a_N}\right)^{1-H}\widehat{V}_N(a_N)\right)_N$ *for* $H = 0.7$*,* $N = 10000$ *and* $a_N = N^{0.6}$ *from* 100 *independent replications.*

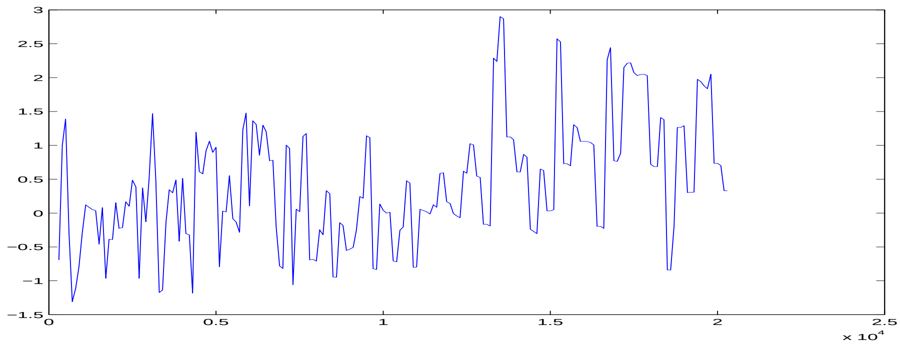

Figure 2: *Convergence of the sequence* $\left(C_{T_2}^{-1}(H)\left(\frac{N}{a_N}\right)^{1-H}\widehat{V}_N(a_N)\right)_N$ *for* $H = 0.7$ *and* $a_N = N^{0.5}$*.*

a lightly smaller variance than it should have in the case $H = 0.9$. Moreover in such a case, there is a very slow convergence rate, *i.e.* $(N/a_N)^{0.1}$, in the limit theorem.

An example of the estimation of the limit density is also presented in Figure 1 in the case $H = 0.7$, $N = 10000$ and $a_N = N^{0.6}$. Such a density is quite similar to a standard Gaussian density but a Kolmogorov-Smirnov test invalides the hypothesis that this distribution is a $\mathcal{N}(0,1)$ law. This result should be compared with the numerical simulation of the Rosenblatt density given in [31].
Finally, Figure 2 shows the convergence of the sequence $\left(C_{T_2}^{-1}(H)\left(\frac{N}{a_N}\right)^{1-H}\widehat{V}_N(a_N)\right)_N$ when N increases. This sequence does not seem to converge in $\mathbb{L}^2(\Omega)$ as it is claimed in Section 3.

5.2. Estimation of H

Here we consider that a sample $(X_1,\ldots,X_N)$ of $X = \sigma^2\, X^H$ with X^H a fBm or a Rosenblatt process is known and H and $\sigma^2 > 0$ are unknown. Denote the sample variance of wavelet coefficients

$$\widehat{I}_N(a_N) := \frac{1}{N_{a_N}}\sum_{j=1}^{N_{a_N}} e^2(a_N,j).$$

Then, following Proposition 3, one deduces that

$$N_{a_N}^{\alpha}\left(\frac{1}{\sigma^2\, C_\psi(H)}\frac{\widehat{I}_N(ia_N)}{(ia_N)^{2H+1}}-1\right)_{1\le i\le \ell}\ \xrightarrow[N\to\infty]{\mathcal{D}}\ (\varepsilon_i)_{1\le i\le \ell}, \tag{35}$$

where the asymptotic distribution $\big(\varepsilon_i\big)_{1\le i\le \ell}$ is a Gaussian distribution (with $\alpha = 1/2$ when X^H is a fBm and $Q \geq 2$) or a Rosenblatt distribution as defined in Theorem 4 (with $\alpha = 1-H$ when X^H is a Rosenblatt process). Therefore, from the so-called Delta-Method, we also have

$$N_{a_N}^{\alpha}\,\Big(\log\big(\widehat{I}_N(ia_N)\big) - (2H+1)\log(ia_N) - \log(\sigma^2\,C_\psi(H))\Big)_{1\le i\le \ell} \xrightarrow[N\to\infty]{\mathcal{D}} \big(\varepsilon_i\big)_{1\le i\le \ell}. \tag{36}$$

Therefore, a log-log-regression of $\Big(\widehat{I}_N(ia_N)\Big)_{1\le i\le \ell}$ by $\big(ia_N\big)_{1\le i\le \ell}$ provides an estimator of H (this estimation method has been introduced in [15]). Such an estimator is defined by

$$\widehat{H}_N := \big(\frac{1}{2}\,,\,0\big)'\cdot\big(Z'_\ell\,Z_\ell\big)^{-1}Z'_\ell\,\big(\log(\widehat{I}_N(ia_N))\big)_{1\le i\le \ell} - \frac{1}{2}, \tag{37}$$

where $Z_\ell(i,1) = \log i$ and $Z_\ell(i,2) = 1$ for all $i = 1,\cdots,\ell$. Then Proposition 3 implies

Proposition 4. *Assume that $\psi \in \mathcal{C}^m(\mathbb{R})$ with $m \geq 1$ and ψ is $[0,1]$-supported. Let $(a_k)_{k\in\mathbf{N}}$ be a sequence of integer numbers satisfying $N\,a_N^{-1} \underset{N\to\infty}{\longrightarrow} \infty$ and $a_N \underset{N\to\infty}{\longrightarrow} \infty$. Let $(X_1,\ldots,X_N)$ an observed sample of $X = \sigma^2\,X^H$ where X^H is a fBm or a Rosenblatt process. Then,*

1. *if X^H is a fBm with $0 < H < 1$ and $Q \geq 2$ or with $0 < H < 3/4$ and $Q = 1$, and if $N\,a_N^{-1-(1\wedge\frac{2m}{3})} \underset{N\to\infty}{\longrightarrow} 0$, then with $\gamma^2(H,\psi,\ell) > 0$ defined in (40) and depending only on H, ψ and ℓ,*

$$\sqrt{\frac{N}{a_N}}\,\big(\widehat{H}_N - H\big) \xrightarrow[N\to\infty]{\mathcal{D}} \mathcal{N}\big(0\,,\,\gamma^2(H,\psi,\ell)\big); \tag{38}$$

2. *if X^H is a Rosenblatt process with $1/2 < H < 1$, $Q \geq 1$ and if $N\,a_N^{-2} \underset{N\to\infty}{\longrightarrow} 0$, then*

$$\big(\frac{N}{a_N}\big)^{1-H}\big(\widehat{H}_N - H\big) \xrightarrow[N\to\infty]{\mathcal{D}} L_{H,\psi,\ell} \tag{39}$$

where $L_{H,\psi,\ell}$ defined in (41) is a distribution depending only on H, ψ and ℓ.

Proof of Proposition 4: From the results of Proposition 3, the relation (35) is clear since $\widehat{I}_N(a) = \widehat{V}_N(a) + 1$ for all $a > 0$. Now using the usual multidimensional Delta-method (see for instance [34]) with the transformation function $g(x_1,\ldots,x_\ell) = \big(\log(x_1),\ldots,\log(x_\ell)\big)'$ applied to the limit theorem (35), one obtains

$$N_{a_N}^{\alpha}\Big(g\Big(\frac{1}{\sigma^2\,C_\psi(H)}\Big(\frac{\widehat{I}_N(ia_N)}{(ia_N)^{2H+1}}\Big)_{1\le i\le \ell}\Big) - g(1,\ldots,1)\Big) \xrightarrow[N\to\infty]{\mathcal{D}} J_g(1,\ldots,1)\,\big(\varepsilon_i\big)_{1\le i\le \ell},$$

where $J_g(1,\ldots,1)$ is the Jacobian matrix of g at point $(1,\ldots,1)$. Therefore, since $J_g(1,\ldots,1)$ is the identity matrix, one obtains (36). Then, $2H+1$ can be estimated from an ordinary least square regression and we obtain:

$$2\,\widehat{H}_N + 1 = \big(1\,,\,0\big)'\cdot\big(Z'_\ell\,Z_\ell\big)^{-1}Z'_\ell\,\big(\log(\widehat{I}_N(ia_N))\big)_{1\le i\le \ell} = M_\ell\,\big(\log(\widehat{I}_N(ia_N))\big)_{1\le i\le \ell},$$

with M_ℓ a $(1\times\ell)$ matrix. Hence, the formula (37) can be deduced. Moreover, (36) implies

$$N_{a_N}^{\alpha}\,\Big((2\,\widehat{H}_N+1) - (2H+1)\Big)_{1\le i\le \ell} \xrightarrow[N\to\infty]{\mathcal{D}} M_\ell\,\big(\varepsilon_i\big)_{1\le i\le \ell}.$$

Therefore, when X^H is a fBm, since $\big(\varepsilon_i\big)_{1\le i\le \ell} \overset{\mathcal{D}}{\sim} \mathcal{N}\big(0\,,\,L_1^{(\ell)}(H)\big)$ from (16), we deduce (38) with

$$\gamma^2(H,\psi,\ell) := \frac{1}{4}\,M_\ell\,L_1^{(\ell)}(H)\,M'_\ell. \tag{40}$$

The same trick implies (39) when X^H is a Rosenblatt process and

$$L_{H,\psi,\ell} := \frac{1}{2}\,M_\ell\,\big(R^H_{1,1},\ldots,R^H_{1,\ell}\big)' \tag{41}$$

with the distribution of $\big(R^H_{1,1},\ldots,R^H_{1,\ell}\big)$ defined in Theorem 4. ∎

Remark 5. *An estimator of σ^2 can also be provided by this method, with the same convergence rate.*

Remark 6. *In Proposition 4 we do not study the case where X^H is a fBm, $Q = 1$ and $3/4 < H < 1$. Indeed, in a statistical framework devoted to the estimation of H, since the choice of ψ is arbitrary, there is no reason to chose ψ with $Q = 1$ which gives a worst convergence rate than ψ with $Q \geq 2$.*

We summarize our results concerning the convergence rate, with $m \geq 2$ and $a_N = N^{1/2+\delta}$ with $\delta > 0$ arbitrary small:

1. if X^H is a fBm and $Q \geq 2$ or $Q = 1$ and $0 < H < 3/4$, the convergence rate of $\widehat{H}_N$ is $N^{1/4-\delta/2}$;
2. if X^H is a Rosenblatt process and $Q \geq 1$, the convergence rate of $\widehat{H}_N$ is $N^{(1-H)/2-\delta(1-H)}$.

Such convergence rates are weak in a parametric framework. For instance, applied to the increments of a fBm, the convergence rate of the maximum likelihood or the approximated Whittle maximum likelihood estimator are $N^{1/2}$ (see [11] or [16]); as far as we know, there are not such results in the case of the Rosenblatt process. An estimator based on quadratic variations has been studied in [33] and a noncentral limit theorem is proved with convergence rate N^{1-H}; but such an estimator is almost parametric and can not be applied to processes which are not strictly self-similar. As it was previously recalled in the introduction, the wavelet based estimator is interesting because it can also be applied in a semiparametric framework.

Remark 7. *In a statistical framework when X^H is a Rosenblatt process, $(X_1, \cdots, X_N)$ is known but H unknown, how to obtain confidence interval from (39)? We propose a three step procedure. The first step is to chose $a_N = [N^{1/2+\delta}]$ with $\delta > 0$ arbitrary small (for instance $\delta = 0.05$) and compute $\widehat{H}_N$. The second step consists in the computation of the quantile q (for instance the 97.5%-quantile) of the distribution of $L_{\widehat{H}_N,\psi,\ell}$ (defined by (39)) from Monte-Carlo simulations of Rosenblatt processes with parameter $H = \widehat{H}_N$. Indeed, it is possible to replace (39) by:*

$$\Big(\frac{N}{a_N}\Big)^{1-\widehat{H}_N}\big(\widehat{H}_N - H\big) - L_{\widehat{H}_N,\psi,\ell} \xrightarrow[N\to\infty]{\mathcal{D}} 0,$$

because $L_{\widehat{H}_N,\psi,\ell} - L_{H,\psi,\ell} \xrightarrow[N_a\to\infty]{\mathrm{P}} 0$ and $\big(\frac{N}{a_N}\big)^{1-\widehat{H}_N} \times \big(\frac{N}{a_N}\big)^{H-1} = \exp\big[(H - \widehat{H}_N)\log\big(\frac{N}{a_N}\big)\big] \xrightarrow[N_a\to\infty]{\mathrm{P}} 1$ from (39) and Slutski's Lemma. Finally, the confidence interval of H will be $\big[-q\,\big(\frac{N}{a_N}\big)^{\widehat{H}_N-1}, q\,\big(\frac{N}{a_N}\big)^{\widehat{H}_N-1}\big]$ since the density of $L_{\widehat{H}_N,\psi,\ell}$ is clearly an even function.

Since there are not other asymptotic results than ours on semiparametric estimators of H for a Rosenblatt process, we can only numerically compare the wavelet estimator with other estimators. We have chosen two usual semiparametric estimators of long memory parameters: the local Whittle estimator (denoted $\widehat{H}_{Whittle}$) defined in [26] and the adaptive global log-periodogram estimator (denoted $\widehat{H}_{LogPer}$) defined in [22]. Remark that such estimators are applied to the increments of X. Note also that both the Matlbab softwares of these estimation procedures can be downloaded from `http://samm.univ-paris1.fr/-Jean-Marc-Bardet`. For completing the numerical study, we have considered two cases of wavelet based estimators:

- $\widehat{H}_{Haar}$ where ψ is the Haar mother wavelet defined by $\psi(x) = 1$ for $0 < x < 1/2$ and $\psi(x) = -1$ for $1/2 < x < 1$ and elsewhere $\psi = 0$; moreover, we consider $a_N = [N^{0.5}]$ and $\ell = [N^{0.3}]$.
- $\widehat{H}_{\psi_C}$ where ψ_C is wavelet function defined above; moreover, we consider $a_N = [N^{0.4}]$ and $\ell = [N^{0.3}]$.

For both these estimators the number ℓ of scales is chosen as a sequence depending on N because additional simulations have shown that the variance of these estimators is optimal with such a choice. The results of simulations are given in Table 5.2.

The conclusions from Table 5.2 are the following:

- The local Whittle estimator $\widehat{H}_{Whittle}$ is clearly the most accurate. The wavelet based estimator $\widehat{H}_{\psi_C}$ is almost as accurate as $\widehat{H}_{Whittle}$. The estimator $\widehat{H}_{LogPer}$ seems to be a little less accurate while $\widehat{H}_{Haar}$ is not satisfying (it can be explained by the fact that this wavelet function is not continuous in 1/2).

H		0.6		0.7		0.8		0.9	
		Mean	$\sqrt{\text{MSE}}$	Mean	$\sqrt{\text{MSE}}$	Mean	$\sqrt{\text{MSE}}$	Mean	$\sqrt{\text{MSE}}$
$N = 500$	$\widehat{H}_{Haar}$	0.38	0.32	0.43	0.35	0.47	0.37	0.43	0.49
	$\widehat{H}_{\psi_C}$	0.56	0.15	0.62	0.16	0.72	0.15	0.84	0.10
	$\widehat{H}_{Whittle}$	0.62	0.07	0.69	0.07	0.77	0.08	0.87	0.09
	$\widehat{H}_{LogPer}$	0.58	0.17	0.67	0.16	0.71	0.18	0.77	0.24
$N = 2000$	$\widehat{H}_{Haar}$	0.50	0.18	0.60	0.18	0.70	0.16	0.83	0.11
	$\widehat{H}_{\psi_C}$	0.59	0.07	0.65	0.07	0.75	0.08	0.85	0.07
	$\widehat{H}_{Whittle}$	0.65	0.06	0.71	0.04	0.80	0.05	0.88	0.05
	$\widehat{H}_{LogPer}$	0.60	0.08	0.65	0.10	0.75	0.12	0.82	0.13
$N = 10000$	$\widehat{H}_{Haar}$	0.54	0.16	0.58	0.15	0.66	0.16	0.69	0.22
	$\widehat{H}_{\psi_C}$	0.60	0.04	0.66	0.05	0.76	0.05	0.86	0.04
	$\widehat{H}_{Whittle}$	0.64	0.04	0.72	0.03	0.79	0.03	0.89	0.03
	$\widehat{H}_{LogPer}$	0.60	0.05	0.67	0.06	0.74	0.08	0.85	0.07

Table 2: *Empirical mean and $\sqrt{MSE}$ of $\widehat{H}_{Haar}$, $\widehat{H}_{\phi_C}$, $\widehat{H}_{Whittle}$, $\widehat{H}_{LogPer}$ for different choices of H and N, from* 100 *independent replications.*

- It appears that the convergence rate of the estimators depends on N and all the estimators seem to be consistant.
- The variance of $\widehat{H}_{LogPer}$ seems to increase as N^{1-H} when H increases unlike both wavelet estimators while theoretical results say that their variances behave as N^{1-H} (see Proposition 4). An explanation of this phenomenon is the following: the asymptotic variance $\text{Var}\,(\widehat{H}_\psi)$ of $\widehat{H}_{\psi_C}$ or $\widehat{H}_{Haar}$ behaves as the $(2,2)$-component of the matrix $\frac{C^2_{T_2}(H)}{4}\big(Z'_\ell\, Z_\ell\big)^{-1} Z'_\ell\, \Sigma_\ell\, Z_\ell\, \big(Z'_\ell\, Z_\ell\big)^{-1}\Big(\frac{a_N}{N}\Big)^{2-2H}$ where Σ_ℓ is given in (29). After computations we obtain that

$$\begin{aligned}\text{Var}\,(\widehat{H}_\psi) &= \frac{C^2_{T_2}(H)}{4}\Big(\frac{\sum_{i=1}^\ell \log i \times \sum_{i=1}^\ell i^{1-H} - \ell \sum_{i=1}^\ell i^{1-H}\log i}{\ell \sum_{i=1}^\ell \log^2 i - \big(\sum_{i=1}^\ell \log i\big)^2}\Big)^2\Big(\frac{a_N}{N}\Big)^{2-2H} \\ &\sim \frac{(1-H)^2\, C^2_{T_2}(H)}{4\,H^4}\,\ell^{2H-2}\Big(\frac{a_N}{N}\Big)^{2-2H} \qquad (\ell\to\infty).\end{aligned}$$

Therefore the larger ℓ the smaller $\text{Var}\,(\widehat{H}_\psi)$. Moreover if $N = 10000$, $a_N = [N^{0.4}]$ and $\ell = [N^{0.3}]$, with $\psi = \psi_C$, we obtain for $\text{Var}\,(\widehat{H}_\psi) \simeq (0.024)^2$, $(0.028)^2$, $(0.030)^2$ and $\simeq (0.030)^2$ for respectively $H = 0.6$, 0.7, 0.8 and 0.9. It explains why the different $\sqrt{\text{MSE}}$ do not numerically seem to depend on H even if they theoretically depends on H.

References

[1] P. Abry, P. Flandrin, M.S. Taqqu and D. Veitch, D. (2003), Self-similarity and long-range dependence through the wavelet lens, In P. Doukhan, G. Oppenheim and M.S. Taqqu editors, Long-range Dependence: Theory and Applications, Birkhäuser.

[2] P. Abry and V. Pipiras (2006), Wavelet-based synthesis of the Rosenblatt process. Signal Processing, 86, 2326-2339.

[3] J. M. Bardet (2002), Statistical study of the wavelet analysis of fractional Brownian motion. IEEE Trans. Inform. Theory 48, 991-999.

[4] J.-M. Bardet and P.R. Bertrand (2008), A nonparametric estimation of the spectral density of a continuous-time Gaussian process observed at random times. To appear in Scand. J. Statist.

[5] J.M. Bardet, H. Bibi, H. and A. Jouini (2008) Adaptive wavelet based estimator of the memory parameter for stationary Gaussian processes. *Bernoulli*, 14, 691-724.

[6] J-M. Bardet, G. Lang, E. Moulines and P. Soulier (2000), Wavelet estimator of long-range dependent processes. Stat. Inference Stoch. Process. 4, 85-99.

[7] J. Beran (1994), Statistics for Long-Memory Processes. Chapman and Hall.

[8] P. Breuer and P. Major (1983), Central limit theorems for nonlinear functionals of Gaussian fields. J. Multivariate Anal. 13, 425-441.

[9] A. Chronopoulou, C.A. Tudor and F. Viens (2009), Application of Malliavin calculus and analysis on Wiener space to long-memory parameter estimation for non-Gaussian processes. C.R.A.S. Mathématiques, 347, 663-666.

[10] A. Chronopoulou, C.A. Tudor and F. Viens (2009), Variations and Hurst index estimation for a Rosenblatt process using longer filters, Electronic Journal of Statistics, 3, 1393-1435

[11] R. Dahlhaus (1989), Efficient parameter estimation for self-similar processes. Ann. Statist. 17, 1749-1766.

[12] R.L. Dobrushin and P. Major (1979), Non-central limit theorems for non-linear functionals of Gaussian fields. Z. Wahrsch. Verw. Gebiete, 50, 27-52.

[13] P. Doukhan, G. Oppenheim, and M.S. Taqqu (Editors) (2003), Theory and applications of long-range dependence, Birkhäuser.

[14] P. Embrechts and M. Maejima (2002), Selfsimilar processes. Princeton University Press, Princeton, New York.

[15] P. Flandrin (1992), Wavelet analysis and synthesis of fractional Brownian motion. IEEE Trans. on Inform. Theory, 38, 910-917.

[16] R. Fox and M.S. Taqqu (1986). Large-sample properties of parameter estimates for strongly dependent Gaussian time series. Ann. Statist. 14, 517-532.

[17] R. Fox and M.S. Taqqu (1987). Multiple stochastic integrals with dependent integrators. J. Multivariate Anal. 21, 105-127.

[18] L. Giraitis and D. Surgailis (1990). A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotical normality of Whittle's estimate. Probab. Theory Related Fields, 86, 87-104.

[19] L. Giraitis and M.S. Taqqu (1999). Whittle estimator for finite-variance non-Gaussian time series with long memory. Ann. Statist. 27, 178-203.

[20] Y. Meyer, F. Sellan and M.S. Taqqu (1999). Wavelets, generalized white noise and fractional integration: the synthesis of fractional Brownian motion. J. Fourier Anal. Appl. 5, 465-494.

[21] E. Moulines, F. Roueff and M.S. Taqqu (2008), Central Limit Theorem for the log-regression wavelet estimation of the memory parameter in the Gaussian semi-parametric context. Fractals, 15, 301313.

[22] E. Moulines and P. Soulier (2003). Semiparametric spectral estimation for fractional processes. In P. Doukhan, G. Oppenheim et M.S. Taqqu editors. Theory and applications of long-range dependenc, 251–301, Birkhäuser, Boston.

[23] D. Nualart (2006), Malliavin calculus and related topics, 2nd ed. Springer.

[24] D. Nualart and S. Ortiz-Latorre (2008), Central limit theorems for multiple stochastic integrals and Malliavin calculus. Stochastic Process. Appl. 118, 614-628.

[25] D. Nualart and G. Peccati (2005), Central limit theorems for sequences of multiple stochastic integrals. Ann. Probab. 33, 173-193.

[26] P.M. Robinson (1995). Gaussian semiparametric estimation of long range dependence, Ann. Statist., 23, 1630-1661.

[27] F. Roueff and M. S. Taqqu (2009). Asymptotic normality of wavelet estimators of the memory parameter for linear processes. To appear in J. Time Ser. Anal.

[28] G. Samorodnitsky and M. Taqqu (1994), Stable Non-Gaussian random variables. Chapman and Hall, London.

[29] M. Taqqu (1975), Weak convergence to the fractional Brownian motion and to the Rosenblatt process. Z. Wahrsch. Verw. Gebiete 31, 287-302.

[30] M. Taqqu (1978), A representation for self-similar processes. Stochastic Processes and their Applications, 7, 55-64.

[31] S. Torres and C.A. Tudor (2008), Donsker type theorem for the Rosenblatt process and a binary market model. Stoch. Anal. Appli. 27, 555-573.

[32] C.A. Tudor (2008), Analysis of the Rosenblatt process. ESAIM Probability and Statistics, 12, 230-257.

[33] C.A. Tudor and F. Viens (2009), Variations and estimators for the self-similarity order through Malliavin calculus. Ann. Probab. 37, 2093-2134.

[34] Van der Vaart, A. (1998), Asymptotic statistics. Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge.